AN INVESTIGATION OF THE ROLE OF PROPOSITIONAL LOGIC IN PERFORMANCE ON PIAGETIAN FORMAL OPERATIONAL TASKS

BY

SUSAN AILEEN WARNER

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1981

To Aileen E. Warner

and the late

Stephen E. Warner

ACKNOWLEDGEMENTS

I would like to thank Drs. Mary L. Koran, Patricia Miller, Scott Miller, and Keith White for their enthusiasm and assistance on this dissertation and for serving as members of my committee. I would also like to thank Libbie Kirkley-Best for her very competent renderings of stimulus materials. Most of all, I would like to thank my committee chairman, Dr. Richard A. Griggs, for his unfailing enthusiasm and encouragement, without which this dissertation would not have been completed.

TABLE OF CONTENTS

APPENDIX		Page
C	THE PROBLEM BOOKLET USED IN EXPERIMENT 1	111
D	THE PERMUTATION PROBLEM PRESENTED IN EXPERIMENT 1	115
E	THE INFORMATION SHEET AND TWO PROBLEMS USED IN EXPERIMENT 2	116

LIST OF TABLES

Table		Page
1	The 16 Binary Propositional Logic Rules	6
2	The Order of Tasks in Experiment 1	18
3	Mean Number, Standard Deviation, and Range of Classes Taken by Subjects in Experiment 1	50
4	The Scoring System for the Three Traditional Piagetian tasks in Experiment 1	52
5	The Number of Subjects at Each Level and Mean Scores and Standard Deviations for the Three Traditional Piagetian Problems Given in Experiment 1	53
6	The Response Patterns for the Selection Task in Experiment 1	58
7	The Response Patterns for the THOG Problem in Experiment 1	61
8	The Correlation Matrix for the Factors in Experiment 1	63
9	Significant Partial Correlations between Selected Factors in Experiment 1	68
10	The Best One-, Two-, and Three-Variable Models for Selected Factors in Experiment 1	71
11	The Number of Subjects for Each Type of Answer Given to Question 8 on the Bending Rods Problem (Experiment 2)	87
12	The Number of Subjects for Each Type of Answer Given to Question 9 on the Bending Rods Problem (Experiment 2)	89
13	The Correlation Matrix for the Factors in Experiment 2	95

Abstract of Dissertation Presented to the
Graduate Council of the University of Florida
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

AN INVESTIGATION OF THE ROLE OF PROPOSITIONAL LOGIC IN PERFORMANCE ON PIAGETIAN FORMAL OPERATIONAL TASKS

By

Susan Aileen Warner

December 1981

Chairperson: Richard A. Griggs
Major Department: Psychology

Piaget hypothesized that the abilities characterizing formal reasoning are the result of a cognitive structure consisting of the 16 binary operations of propositional logic. In Experiment 1 this hypothesis was investigated by measuring performance on paper-and-pencil versions of four Piagetian tasks (the permutation, combination of chemicals, bending rods, and pendulum problems) and two propositional logic tasks--the selection task (based on the implication rule) and the THOG problem (based on the exclusive disjunction rule). In addition, possible relationships to classroom experience, Scholastic Aptitude Test (SAT) scores, and performance on the Group Embedded Figures Test were investigated.

No relationship between performance on the Piagetian and logic tasks was observed. The best predictor for performance

on the logic tasks and two of the Piagetian tasks (combination of chemicals and permutations) was SAT scores. However, less than 16% of the variance for the logic tasks and approximately 35% of the variance for the Piagetian tasks was accounted for by this factor. The other two Piagetian tasks were the best predictors of performance for each other. In addition, the most important ability necessary for successful performance on the Piagetian tasks was the isolation and control of variables. It was concluded that separate abilities may underlie performance on the two types of tasks.

This conclusion was supported by the results of Experiment 2, which employed a more sensitive test of the propositional logic basis of solving Piagetian tasks. The bending rods and pendulum problems from Experiment 1 were modified by formulating specific questions based on the isolation and control of variables. Hypothetical results and hypotheses based on propositional logic rules were also presented for evaluation by subjects. Most subjects recognized confounded results and could isolate and test variables correctly. Additionally, their responses conformed to simple logical rules when evaluating most of the presented results and hypotheses. However, the relationship between performance on the questions based on propositional logic and on the isolation of variables was inconsistent. The results of both experiments are discussed in terms of the use of "schemes" rather than propositional logic as the basis for successful performance on Piagetian tasks.

CHAPTER 1
PIAGET'S THEORY OF FORMAL OPERATIONS

Since the publication of Inhelder and Piaget's 1958 book entitled The Growth of Logical Thinking from Childhood to Adolescence, formal operational abilities have been the subject of much discussion, controversy and empirical investigation. This area of Piaget's theory is probably the least understood, however, because there is proportionally less inquiry into this last stage of cognitive development than into any other Piagetian stage. This chapter will begin with a description of formal operations as it was originally defined by Piaget. Empirical evidence for Piaget's theory will then be presented, and lastly, the current state of formal operations theory and research will be discussed.

Original Formulation of Formal Operations

Piaget's theory of formal operations is most thoroughly defined in Inhelder and Piaget (1958). Several excellent secondary sources serve to clarify and organize the material presented by Piaget, including books by Brainerd (1978), Flavell (1963), and Ginsburg and Opper (1969). The following discussion draws heavily on the information presented in these three books and in Inhelder and Piaget (1958).

Piaget's theory of cognitive development is characterized by the postulation of four stages--the sensorimotor stage, preoperations, concrete operations, and formal operations. As in accordance with any stage theory, the child passes through these stages in this invariant order. The third stage, concrete operations, is reached by most normal American children by age seven. Since formal operations is defined partially by advances made over concrete operational abilities, the limitations of this stage will be discussed in a fair amount of detail.

The major achievement of concrete operations is that the child possesses reversible operations which, among other things, enable him or her to focus on transformations rather than states, become less egocentric, and decenter his or her thought. This ability is shown clearly in several of Piaget's demonstrations, especially those concerning conservation. For example, the child now realizes that there is the same amount of water in a short wide beaker as in a tall, thin beaker when the water is poured from one to the other. A child in the preoperational stage will state that there is more water in the tall, thin beaker, apparently "captured" by the saliency of water level.

There are several important limitations of concrete operations. Firstly, the reversible operations of negation (undoing) and reciprocity (compensation) that allow the child to achieve the abovementioned concrete operational

abilities exist as cognitive operations but are not coordinated into a single, integrated system. Thus, the child cannot correctly solve problems that require the use of both of these operators. These operators are described by Piaget in terms of logical groupings. These groupings serve as a model for the child's cognitive organization, and the integration of these groupings occurs in formal operations.

A second limitation of concrete operations is that the organizing activity of these operations is focused on concrete objects and events in the immediate present. In other words, reality, rather than the potential, is the starting point for cognitive activity. Limited inferences can be made, such as linear transitive inferences or class inclusion inferences, but only if a concrete referent exists. This link to reality leads to a phenomenon known as horizontal decalage, in which conservation does not become evident for all types of matter at the same time. In fact, years intervene between the attainment of conservation of the earliest and latest materials (usually liquid and volume, respectively). This decalage occurs precisely "because his cognitive instruments are insufficiently 'formal,' insufficiently detached and dissociated from the subject matter they bear upon, to permit a content-free, once-for-all structuring" (Flavell, 1963, p. 204).

The transition to formal operations occurs sometime during adolescence. The major general characteristic of

this stage is that reality is no longer the starting point for a problem. Rather, reality becomes just one special case of all possibilities. Now, instead of trial-and-error or random experimentation, the formal operational adolescent solves a problem by trying to imagine all of the possible relationships present in the data and then attempts to discover which of the possible relations do exist. This discovery takes place through a combination of logical analyses and experimentation. Since the adolescent has the potential to imagine all that could occur, he or she has a much greater chance of discovering all that does occur. The adolescent is no longer limited by the specific situation, which is a fundamental reorientation toward approaching cognitive problems over the concrete operational child.

Several other characteristics of formal thought are a result of this reorientation. Thought becomes primarily hypothetico-deductive. To attempt to discover the real among the possible, one must first consider a set of possible hypotheses to be tested. Those which are confirmed by data are then classified as reality. This type of scientific reasoning is possible because the adolescent is now able to systematically isolate and combine variables into all possible combinations, as well as correctly interpret the results of these empirical tests.

In addition, formal operational thought is no longer content-bound, it is propositional. Concrete operational

thinking is defined as a first-order operation--the child organizes objects and events themselves by classification, seriation, and so on. Formal operational thinking consists of second-order operations. The adolescent takes the results of the first-order operations, formulates propositions from them, and then operates further on these propositions by forming logical connections between them.

Generally, formal operations can be characterized by an orientation toward problem solving by the use of isolation and control of variables and combinatorial analyses to organize data. In addition, there is an orientation towards the hypothetical and towards logical justification and proof.

Piaget proposed that these general characteristics of formal operations are the result of an underlying logicomathematical core system. This core system consists of an integrated lattice-group structure, namely, the 16 binary propositional operations. These 16 operations are given in Table 1. The operations are called "binary" because they consist of two propositions, p and q, which can be either true (p,q) or false ($\bar{p}$,$\bar{q}$). Thus, there are four possibilities--(p, $\bar{p}$, q, and $\bar{q}$), which are combined into four elementary propositions: pq, p$\bar{q}$, $\bar{p}$q, and $\bar{p}\bar{q}$. There are 16 unique patterns of truth and falsity for these four propositions. This is demonstrated in Table 1. Each pattern is named and consists of a propositional logic rule. The understanding and use of these rules has been investigated, and

Table 1

The 16 Binary Propositional Logic Rules

Rule Name[a]	Elementary Propositions[b]			
	pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$
Affirmation	T	T	T	T
Inclusive disjunction	T	T	T	F
Reverse implication	T	T	F	T
Affirmation of p	T	T	F	F
Implication	T	F	T	T
Affirmation of q	T	F	T	F
Biconditional	T	F	F	T
Conjunction	T	F	F	F
Nonconjunction	F	T	T	T
Exclusive disjunction	F	T	T	F
Denial of q	F	T	F	T
Nonimplication	F	T	F	F
Denial of p	F	F	T	T
Nonreverse implication	F	F	T	F
Conjunctive negation	F	F	F	T
Negation	F	F	F	F

[a]These rule names were derived from Brainerd (1978) and Staudenmayer (1975).

[b]A bar over a component proposition represents the false or "not" case.

will be discussed in Chapter 3. For Piaget, each rule represents an operation and is one of the elements of the full integrated lattice structure. Conjunctions are related by "and" (·) and disjunctions are related by "or" (v).

This core system develops from the groupings that characterize concrete operational thinking. The groupings become integrated into a whole structure as a result of increasing disequilibrium because of the inability to successfully deal with problems of increasing complexity.

Piaget provides demonstrations of the use of the 16 binary propositional operations in formal operations throughout Inhelder and Piaget (1958). Fifteen different problems involving different types of experimental apparatus are presented to subjects to illustrate the differences between concrete and formal operational approaches to solving each problem. Not only are the general characteristics of both types of thinking illustrated, but Piaget interprets the formal operational behavior in terms of the 16 binary propositional operations. One example, discussed by Flavell (1963), concerns the combination of colorless chemicals task. In this task, subjects are presented with supplies of four perceptually identical chemicals (colorless and odorless), an indicator chemical (g), and supplies of test tubes. The experimenter shows the subject two beakers which each contain one or more of the four chemicals, and adds g to both. One of the combinations turns yellow and the other does not.

The subject's task is to find out which chemical or combination of chemicals turns yellow when g is added.

First, Piaget assumes that p represents one of the aspects of the problem and q stands for another. For example, in this problem p could represent the presence of a yellow color, $\bar{p}$ its absence, q the presence of one of the chemicals (e.g., chemical A), and $\bar{q}$ its absence. As the adolescent makes the combinations of chemicals, he or she may notice that yellow does appear when chemical A is present (the pq case exists). However, rather than immediately deciding that chemical A causes the yellow color, the adolescent is aware that this reaction is only one of a number of possibilities and that all possible combinations must still be tested. Only when the entire set of combinations in which a yellow color appears or does not appear is known, will the significance of chemical A be known. If $p\bar{q}$, $\bar{p}q$ and $\bar{p}\bar{q}$ also exist, the adolescent will conclude that chemical A is irrelevant to the yellow color (Rule #1 in Table 1). However, if chemical A does cause the yellow color, pq, $p\bar{q}$ and $\bar{p}\bar{q}$ will exist.

A further assertion of Piaget concerning the underlying core structure of thought in formal operations is that the elements of the lattice are integrated by the four operations of Identity (I), Negation (N), Reciprocal (R), and Correlative (C) (the INRC group). These operations allow a given operation to be transformed into a different operation. I, the identity operation, changes nothing in the proposition

on which it is performed. N changes everything in the proposition on which it is performed--positives become negatives, and all conjunctions become disjunctions and vice versa. R changes positives to negatives and vice versa, but leaves conjunctions and disjunctions unchanged. C changes conjunctions to disjunctions and vice versa, but leaves positives and negatives unchanged. Each operation in the lattice structure is related to other operations as its I, N, R, or C operation. Hart (1971) and Brainerd (1978) provide tables which show the specific relationships between the operations. These relationships between the operations are what allow the adolescent to understand reversibility and to form hypotheses based on the results of all possible combinations of variables.

The INRC group can refer to both physical actions and logical operations. As mentioned before, Piaget interprets physical actions in terms of propositions and logical operations to link them to the core structure consisting of the 16 binary propositions. Objections to this use of propositional logic, along with other objections to Piaget's theory, will be discussed in Chapter 2.

Piaget's methodology, as described in Inhelder and Piaget (1958), will now be discussed because of relevance to later experiments. Piaget investigated formal operational thought by means of presenting subjects with a situation, scientific

in nature, and asking them to discover the relevant variables determining the action of interest. His method was a modified-clinical method in which open-ended instructions were given at the beginning of the session, and the actions of the child determined whether further prompts, questions, or explanations would be necessary. Each experimental setting was designed to test scientific reasoning, and some problems focused on specific aspects of formal operations such as isolation of variables or exclusion of variables.

The apparatus was often quite elaborate, as in the combination of colorless chemicals task, which involved five chemicals (water, diluted sulphuric acid, oxygenated water, sodium thiosulphate, and potassium iodide) plus test tubes and beakers. Some of the other tasks are as follows.

The Oscillation of a Pendulum--subjects are presented with lengths of string, weights, and an overhanging bar; and the problem is to find the factor that determines the frequency of oscillation.

The Bending Rods Problem--the subjects is presented with a set of rods differing in composition, length, thickness, and cross-sectional form and a large basin of water. The rods can be attached to the edge of the basin, and three different weights can be screwed onto the end of the rods. The subject is asked to determine if the rod is flexible enough to touch the water.

The Equality of Angles of Incidence and Reflection--the apparatus resembles a billiard game. Balls are launched with a spring device that can be pivoted around a fixed point. The subject is asked to aim at targets placed at different points by shooting the ball against a projection wall so that it rebounds to the interior of the apparatus.

The Law of Floating Bodies--the subject is asked to classify objects as to whether or not they will float on water and to explain their answers. Then, he or she actually experiments with buckets of water and objects including pieces of wood, wire, a metal weight, metal needles, a pebble, a plank, boats, ducks, a piece of candle, matches, cork, paper, and so on.

Other experimental situations include predicting displacement of water and other liquids in communicating vessels, determining why a metal bar attached to a rotating disk stops with the bar pointing to one pair of boxes instead of any other boxes placed around the disk, predicting how far balls of varying sizes will travel when launched by a spring device, predicting in what order different sized balls in various positions will leave a rotating disk when the speed of rotation is increased, and producing two shadows of the same size on a screen using different sized rings in front of a light source.

Empirical Evidence

From the earliest reviews of and replications of the work in Inhelder and Piaget (1958) there have been numerous questions concerning Piaget's interpretation of his data, the use of propositional logic as his model for thought, the universality of formal operations, and the extent to which formal operations is context-free. The most serious criticisms of formal operations such as his use of propositional logic will be discussed in Chapter 2. However, many researchers accept formal operations as a valid psychological construct, and their work has served to refine and expand the theory. A review of representative research in various areas of formal operations follows.

Lovell (1961) was one of the first researchers to attempt to replicate the experiments in Inhelder and Piaget (1958). Using 10 of the tasks and 200 subjects, Lovell interpreted his data as supportive of Piaget in stages of development, the overall level of performance on all tasks, and the type of reasoning used as evidenced by the subject's protocols. However, there were enough discrepancies in the amount of prompting required, the order of behaviors evidenced, and other characteristics that Lovell concluded that Inhelder and Piaget "have somewhat forced the development of the child's thinking into a theoretical framework" (Lovell, 1961, p. 151).

Martorano (1977) also administered 10 traditional Piagetian tasks to 20 female subjects each in grades 6, 8, 10, and 12. Results revealed that as grade increased, mean scores on the 10 tasks increased as well, with evidence that formal thinking begins to emerge between 12 and 15 years of age. However, subjects even in the oldest age group did not perform consistently at the formal operational level across all tasks. Post hoc analyses revealed significant differences between tasks in level of performance. The order of difficulty of the tasks, from easiest to hardest, was colored tokens, correlations, combination of colorless chemicals, permutations, pendulum, bending rods, shadows, balance, communicating vessels, and hydraulic press. Other researchers have avoided the effects of the task differences by testing primarily for only one of the aspects of formal operational ability, such as isolation of variables (e.g., Kuhn & Angelev, 1976, Kuhn & Brannock, 1977), combinatorial reasoning (Barrett, 1975), and proportional reasoning, (e.g., Capon & Kuhn, 1979).

It soon became evident that the attainment of formal operations was far from universal. A sample of estimates of the number of subjects in high school and in college operating in a fully formal operational manner range from 20% to 25% (e.g., McKinnon & Renner, 1971, Schwebel, 1975; Tomlinson-Keasey, 1972). Higher estimates occur for subjects operating at a lower formal operational stage, usually

around 50% (e.g., Danner & Day, 1977; Martorano, 1977; Tomlinson-Keasey, 1972). In addition, there is some evidence for sex differences in the attainment of formal operations favoring males (e.g., Douglas & Wong, 1977, Keating & Schaefer, 1975, Schwebel, 1975). Piaget (1972) clarified his theory to take into account specific knowledge, professions, and interests which may have an effect on the expression of formal operations. He argued that this was not contradictory to his earlier assertion that formal thinking is independent of content because ". . . it is one thing to dissociate the form from the content in a field which is of interest to the subject and within which he can apply his curiosity and initiative, and it is another to be able to generalize this same spontaneity of research and comprehension to a field foreign to the subject's career and interests" (Piaget, 1972, p. 10). Piaget (1972), then, argues that formal operations are free from concrete content but only when the situation involves vital interests or knowledge. This may explain both the scarcity of formal operational behavior using the traditional science-based tasks and the observed sex differences, because a minority of students may be interested in science, especially female students.

Another explanation for the lack of formal operational performance for many subjects is the method of assessment used. Traditional formal operational tasks use open-ended instructions, and while this may allow for a wider range of

abilities to be expressed, it may not be clear to some subjects what the optimal procedure for successful performance may be. Danner and Day (1977) tested this hypothesis by providing a series of prompts (from less to more explicit) until the subject made unconfounded tests of all variables or these tests were demonstrated (the last prompt). Older subjects (aged 17 years) received less prompts but demonstrated improved performance over tasks--95% of these subjects used a formal operational strategy for a third unprompted task. A control group who did not receive prompts did not improve, so practice was not the reason for improvement. Rather, the prompts served to clarify the tasks for the subject.

Another study hypothesized that the lack of formal operational behavior for many subjects may be the result of a variety of superficial cognitive difficulties such as verbal articulation and organizational difficulties. Kuhn, Ho and Adams (1979) gave college students and fourth and fifth graders who had been classified as non-formal operational repeated exposure to problems requiring formal reasoning. For 20 to 30 minutes per week for 11 weeks, subjects were asked questions about a chemical combinations problem that encouraged them to analyze and attempt to interpret what was being observed. No reinforcements, solutions, or strategies for obtaining a solution were suggested. The pre-adolescents made only gradual modest gains, while most of the college

students showed immediate and substantial formal reasoning.

In accordance with Piaget's (1972) hypothesis, several studies have also investigated the role of the content of the tasks on successful formal operational performance, but with mixed results. Bart (1972) concluded that the type of premise (absurd or abstract) in a reasoning task was less important for performance at various levels of formal reasoning than the logical structure of premises. Kuhn and Brannock (1977) devised a task which tests primarily isolation of variables in a more natural setting--the care and health of plants. The same errors of reasoning present before the acquisition of formal operations described by Inhelder and Piaget (1958) occurred with this task, but performance on this task and a traditional Piagetian task (the pendulum problem) was not related. Kuhn and Brannock conclude that the two types of situations pose different problems for subjects.

Capon and Kuhn (1979) tested the use of propositional reasoning by adult female shoppers in a supermarket. The subjects were asked to determine which of two sizes of garlic powder or deodorant was the better buy. Even when the ratio of sizes was simple (2:3), only 32% of the subjects were able to use a proportional reasoning strategy to determine the correct choice. Additionally, Keating and Clark (1980) found that performance on social, interpersonal

reasoning tasks (presumably more familiar to adolescents) was no better than on tasks based in the physical domain. Only 13% of the subjects performed better on the social tasks and 52% of the subjects performed better on the physical tasks. However, Sabatini (1980) found that there were no age effects on formal operational performance between college-aged students and older adults. All of the differences observed were accounted for entirely by professional specialization (scientists versus nonscientists). Also, Pulos and Linn (in press) found familiarity effects on the use of controlling variables. Rural adolescents performed better than urban adolescents on a task involving fishing, and urban subjects performed better on a more scientific task involving a spring apparatus.

A related question concerns the relationship between formal operational abilities and classroom training. Most of the empirical results appear to be negative, despite Piaget's (1972) hypothesis of the effects of familiarity of materials on formal reasoning. Lovell (1961) found no effect of classroom training. Lawson and Blake (1976) tested high school biology students on biology-based formal operational tasks, traditional Piagetian tasks (physics-based), and formal operational tasks with a nonscience content, but found no differences in performance for the three types of tasks. Griffiths (1974) observed that experience in highly technical physics classes was not necessary for formal operational

development; and in fact, many students had technical vocabularies but did not fully possess the formal operational abilities to isolate and control variables on a physics task. On the other hand, Pulos and Linn (in press) and Sabatini (1980), as mentioned above, have found evidence for the effects of experience on formal operational reasoning.

Researchers have also investigated the relationship between formal operational abilities and other variables. Wyatt and Geis (1978) observed that the highest-level formal operational subjects had a greater amount of recall and recall organization than did lower-level formal operational subjects on a multitrial free-recall task. Berzonsky, Weiner and Raphael (1975) hypothesized that subjects experiencing an identity crisis, in which a person considers what one could possibly become in addition to what one is, would perform better on formal reasoning measures than noncrisis subjects. This hypothesis was not supported, and the results were interpreted as in accordance with Piaget's (1972) hypothesis for the selectiveness of formal operations. Blasi and Hoeffel (1974) also agree that personality development in adolescence takes place independently of formal operations because nonphysical concepts such as will and creativity are a part of the integration of concepts of socio-psychological reality.

Neimark (1975; Note 1) has argued for the existence of a relationship between formal operations and field dependence/

independence both empirically and theoretically, as has Pascual-Leone (e.g., 1970) and his colleagues. They propose that field independent subjects are able to abstract out the important attributes of the situation at hand to aid them in successful performance, especially in an ambiguous situation, as traditional formal operational tasks often are (Neimark, Note 1). Flexer and Roberge (1980) have produced evidence that the effects of field dependence/independence are due to an overlap with IQ. Other studies relevant to the relationship of IQ and formal operations have demonstrated that measures of IQ and reasoning abilities tap different abilities (Stephens, McLaughlin, Miller & Glass, 1972), that IQ is related to the speed of mastery of within-stage skills (Webb, 1974), that high IQ students have high levels of formal operational reasoning (Keating, 1975), and that Scholastic Aptitude Test scores are related to formal operational scores for females but not males (Schwebel, 1975).

Arlin (1975, 1977) has proposed that a fifth stage of development exists beyond formal operations (or problem-solving) which is referred to as problem-finding. Her results indicated that formal operations were a necessary but not sufficient condition for the existence of problem-finding. However, Fakouri (1976) argues that problem finding may be viewed as content improvement within formal operations rather than a structural stage, or selective manifestation of formal operational abilities. In addition,

Cropper, Meck and Ash (1977) found no relationship between problem solving and problem finding.

Finally, several researchers have called for an information-processing approach to Piagetian tasks so that attentional, memory and strategy demands on performance on the very complex formal operational tasks can be assessed. Among these neo-Piagetian researchers are Case (e.g., 1978), Pascual-Leone (e.g., 1970), Scatalno (1978), Scardamalia, (e.g., 1977), Siegler (e.g., 1978) and Wollman (e.g., 1979).

To summarize, one unequivocal finding in the investigation of Piaget's theory of formal operations is that the universality of the attainment of formal operations does not appear to be as high as was expected by Inhelder and Piaget. Taking this finding into account, Piaget's results have been replicated with respect to isolation of variables, proportional reasoning, hypothesis testing, and combinatorial analysis, as well as with studies employing at least 10 of his tasks. Conflicting results have been observed for many aspects of formal operations, such as the role of context and the relationship of formal operations to other variables including field dependence/independence, classroom training, and intelligence. The effect of the often-used, open-ended instructions and lack of exposure to the tasks used have been offered as possible reasons for the failure of many subjects to exhibit formal operations as well.

Conflicting results have been reported for the existence of a fifth stage of cognitive development referred to as problem-finding. In addition, a call for task analyses as well as the investigation of the role of information processing abilities in the expression of formal operational abilities for Piaget's tasks has led to the formulation of several neo-Piagetian approaches.

The empirical investigation of the underlying propositional logic basis of formal operations is practically nonexistent. It is not known if this is because researchers assume that this logic base exists and does not need to be empirically verified, if it is assumed that it does not exist and thus does not need to be investigated, or that most researchers simply do not care to do research in this area.

The investigation of the relationship of performance on propositional logic tasks to formal operational ability is the purpose of this dissertation. Hence, the next chapter details some of the theoretical objections to Piaget's propositional logic system, empirical evidence supporting this system, and other criticisms of Piaget's theory. Chapter 3 then discusses the research on two propositional logic problems very similar in overall problem-solving framework to Piaget's problems that will be utilized in this dissertation.

CHAPTER 2
CRITICISMS OF PIAGET'S THEORY OF FORMAL OPERATIONS

Criticisms of Piaget's theory of formal operations stem from several different sources. Much of the confusion about his theory comes from the difficulty in reading Inhelder and Piaget (1958). For example, the procedures for all of the tasks are never fully explained, and the relationship between behavior and Piaget's interpretation of that behavior is not explicit (Keating, 1980). The criticisms discussed in this chapter concern both Piaget's use of propositional logic and some general criticisms of the internal logic of his theory. A reply to the propositional logic criticisms will be presented, as well.

The deficiencies of the logic used by Piaget in describing formal operations are discussed by Brainerd (1978), Ennis (1975), Keating (1980) and Parsons (1960), among others. Ennis (1975) states that Piaget's logic is a "defective" logic because it is not abstract. Piaget's propositions do not stand alone (as abstract propositions do) unless one assumes that there is an explicit existential quantifier attached to the proposition. For example, instead of the proposition "if p then q," Piaget's propositions must be interpreted as "There exists a p such that p is q." This interpretation

must occur because Piaget treats the propositions as sentences, and the variables have concrete referents. Thus, for any propositional statement, all of the combinations mentioned in the statement must exist, and the existence of any unmentioned kind of case is implicitly denied.

Piaget agrees that this is the interpretation that should be given to his logic (Piaget, 1967; cited in Ennis, 1975). However, this conceptualization leads to various inconsistencies as well as paradoxical requirements of the real world, and Piaget does not make it clear how these problems should be resolved. For example, Ennis (1975) argues that it is unreasonable to require the existence of each of the mentioned cases within a rule for that rule to be true. In the example discussed earlier concerning the effects of chemical A in the chemical combination task, the $p\bar{q}$ case must exist (according to the reverse implication rule) if chemical A does cause the yellow color. Thus, a yellow color must exist in the absence of chemical A if chemical A causes the yellow color.

An additional problem concerning the confirmation of the truth of propositions is that Piaget does not make it clear whether subjects must consider all examined cases or all possible cases, including the unexamined cases that are not present in the experimental situation. Sometimes Piaget discusses the actual existence of cases (Inhelder & Piaget,

1958 e.g., p. 299), and at other times he requires the knowledge of the possibility of existence (Inhelder & Piaget, 1958, p. 17).

Ennis points out that Piaget is also vague when he states that children under 11-12 cannot "handle propositional logic." The criteria for handling propositional logic are not clear. Piaget almost certainly does not refer to mastery of propositional calculus because even college students do not find this easy. Moreover, Piaget states that a person using formal operations is mostly unaware of the internal logic structures that he or she is using except for a feeling of certainty (Inhelder & Piaget, 1958). On the other hand, Piaget probably does not wish to state that children under 11 years old cannot reason with any of the forms of propositional logic because evidence against this statement has been known at least since Burt's early experiments (Burt, 1919). In addition, several studies demonstrate that young children are better at making judgments about valid forms of arguments than invalid forms, but the same differentiation holds for adults and adolescents (e.g., O'Brien & Shapiro, 1968; Roberge, 1970; Shapiro & O'Brien, 1970). Thus, Piaget's formal operations theory is challenged on two counts--both children and adolescents have the ability to reason with at least some of the forms of propositional logic and both find it difficult to reason with invalid forms. Also, this implies that the 16 binary propositions

are not operating with equal ease as would be expected if they were connected together in a structured whole, as Piaget postulates (Inhelder & Piaget, 1958).

These difficulties with Piaget's logic system and with interpreting what Piaget actually means in his statements led Ennis to argue that Piaget's claim about formal operational ability appears to be about a total, holistic logical ability and not the ability to reason according to a set of principles of propositional logic. Because of the ambiguities and idiosyncracies present in Piaget's discussions of formal operations, his claims are untestable (Ennis, 1975; Keating, 1980). Piaget's experiments, then, become demonstrations which he interprets within his framework, but the link between the data and his theory is obscure. One well known example is the protocol of GOU which appears to be the only data which demonstrate a subject's ability to state all 16 logical combinations. However, when this protocol was reanalyzed by logicians, evidence for only eight of the 16 binary operations was found (Bynum, Thomas & Weitz, 1972).

This example also demonstrates the difficulty in assessment of formal operations. In several instances, Piaget rejects a verbal criterion. In other words, he does not expect a person with formal operations to be able to use the language of propositional calculus (Inhelder & Piaget, 1958, e.g., p. 254). Yet, he himself uses this criterion exclusively to interpret GOU's protocol (Inhelder & Piaget, 1958,

p. 102). Keating (1980) also discusses problems with the choice of an assessment criterion. Certain tasks that Piaget discusses, such as the combination of chemicals task, suggest that the ability to generate all possible combinations of the elements is a mark of formal operational thought. However, based on his discussion of the integration of the logical system by the INRC group operators, the use of a more stringent criterion of being able to understand and use effectively the total logical system would seem appropriate. For example, generating all possible combinations of a number of elements has become a criterion for the assessment of formal operational thought (e.g., Neimark, 1975).

A second problem that Keating (1980) discusses is the apparent inability to disconfirm the theory. Specifically, as has been mentioned before, it has been observed in many studies that formal operations is not universally attained. Piaget (1972) argues that differential familiarity with the content area probably inhibits performance, and that horizontal decalage may also contribute to a performance failure. Thus, when the theory states that formal operational reasoning should be occurring, the failure of this ability to appear is usually attributed to extraneous task factors. However, where formal operations are presumed not to exist performance failures are attributed to the absence of formal operations.

Another source of difficulty is the apparently multiple usage of terms in Inhelder and Piaget (1958). Blasi and Hoeffel (1974) state that Piaget uses the term reflective thinking in at least three different ways but never distinguishes between these uses. The most common interpretation of reflective thinking is the use of an elementary process by a more complex mechanism, but this term is also used to refer to the knowledge of an act of thinking and to the knowledge of the self. Flavell (1963) points out two uses of the INRC group that Piaget does not distinguish--a physical INRC group in which actual physical transformations occur, and a logical INRC group in which transformation of propositions occur. It is to this multiple usage that Ennis (1975) addresses his critique.

The argument against Piaget's use of propositional logic is similar to a very basic argument against using any type of formal logic as a model of reasoning. Osherson (1975) and Braine (1978) have argued that standard logic cannot mirror human reasoning because logical reality does not correspond to a psychological reality. Paradoxes present in logic exist because certain rules of logic do not correspond to our intuitions about not only the meaning of "if-then" statements, but of "all-are" statements as well. Thus, this argument extends to include all forms of standard logic.

In addition, the translation of any logical calculus into linguistic form will lead to a certain amount of

imprecision. The English language is not as precise as logical calculus, so several different logical interpretations can be made for one statement. Also, Bynum, Thomas and Weitz (1972) state that there are six propositional logical rules which cannot be translated into the English language.

Since Piaget does not distinguish between logical operations on physical objects and on propositions, Ennis' criticisms of his use of logic are accurate but unimportant to Piaget from a psychological standpoint. The empirical investigation of many logic systems has led to a reference to concrete objects for the fullest understanding of subjects' usage of the logic system. Therefore, any psychological investigation into a logic system will involve existential referents to some extent. Ennis' (1975) criticisms appear to be addressed to an inevitable byproduct of the psychological investigation of logic (see Broughton, 1981, for a similar argument). His criticisms are theoretical, not empirical, and only a few empirical studies have been performed relevant to this question.

Piaget's system of interpretation of behavior in terms of his logical system is not clear, as evidenced by the reanalysis of GOU's protocol by Bynum, Thomas and Weitz (1972). The status of the underlying cognitive structure as consisting of the 16 binary propositional operations, therefore, is equivocal. Most researchers in the area of formal operations have not addressed this question. Two studies designed

to investigate this specific question were interpreted as supportive of Piaget's position, and are presented below.

Moshman (1977) tested the knowledge of implication and disjunction with colors and shapes. Subjects chose patterns consistent with a statement using a six-alternative forced choice paradigm. Task difficulty was manipulated by negating the antecedent. Disjunction proved to be easier than implication, but there was evidence for converging development of the knowledge of these rules. Disjunction knowledge begins earlier but the use of the two rules reaches full consolidation with respect to negation at approximately the same time. Moshman maintains that this is the expected result if the two operations are part of a larger <u>structure d' ensemble</u>, as postulated by Inhelder and Piaget (1958).

Riley (1979) presented subjects with a list of 16 propositions, and asked them to choose and record which of four cards (representing the pq, $p\bar{q}$, $\bar{p}q$, and $\bar{p}\bar{q}$ cases) satisfied the conditions in the proposition. He then subjected the results to a Guttman scaling analysis. Results revealed that the single-set propositions were logically prerequisite to the double-set operations, which were prerequisite to the triple-set operations. Riley interpreted these results as supportive of Piaget's contention that the four elementary propositions form the basis of the lattice structure.

Thus, there seems to be some evidence for Piaget's hypothesis concerning the structure of the proposed system of

the 16 binary propositional operations, but the relationship of this structure to behavior is still unclear. Research investigating performance on propositional rules themselves could clarify this issue although the goal of this research was not to test assumptions about formal operations. A review of two particularly relevant propositional logic tasks, Wason's selection task and THOG problem, will be presented in Chapter 3.

CHAPTER 3
PROPOSITIONAL LOGIC TASKS

The understanding of propositional logic has been investigated for many years. The typical deductive reasoning study has asked directly about logical syllogisms, however, by presenting premises and requiring subjects to produce a conclusion, verify a conclusion, or choose the correct conclusion from several alternatives (e.g., Neimark & Slotnick, 1970; Paris, 1973; Roberge, 1971; Rips & Marcus, 1977, Staudenmayer & Bourne, 1977; Wildman & Fletcher, 1977). In addition, inductive reasoning studies employing concept-learning paradigms have been based on propositional rules (e.g., Bourne, 1970; Bruner, Goodnow & Austin, 1956). Because of the vast differences in task framework between these types of tasks and the traditional Piagetian tasks, the results from these investigations may not be relevant to the questions posed in the previous chapter. However, two problems devised by Peter Wason are ideal for the investigation of formal operational abilities because a problem solving framework is utilized for such task. The subject is not asked directly about the rule upon which the problem is based, but if the problem is to be solved correctly, the answers must conform to the logical structure of the rule. Hypothesis-testing ability is also an important aspect of the ability

to solve the problems. These two tasks are discussed in detail below.

Wason's Selection Task

In 1966, Wason reported a problem which is based on the propositional logic rule of implication. This problem is known as the "four-card problem" or the "selection task." The basic task is as follows. The subject is presented with four cards showing, for example, A, D, 4 and 7, respectively. The subject is told that each card has a letter on one side and a number on the other side. The following rule is also presented--"If there is a vowel on one side of a card, then there is an even number on the other side." The task is to select only those cards which are necessary to turn over to determine if the rule is violated.

Since the rule is an implication rule ("if p then q") and the array of four cards represent the p, $\bar{p}$, q and $\bar{q}$ cases, respectively, the correct answer is to turn over the p and $\bar{q}$ cards (A and 7 in the example). Only the combination of p and $\bar{q}$ can falsify the implication rule (see Table 1, Rule 5) so the p card must be checked to see if $\bar{q}$ is on the reverse side, and the $\bar{q}$ card must be checked to see if p is on the reverse side.

This task is extremely difficult for most subjects--usually less than 10% of the subjects product the correct

solution (see Wason & Johnson Laird (1972) and Evans (1978) for reviews of the literature on this task). The two most common errors that are observed are (1) a failure to select the $\bar{q}$-card, and (2) selecting the q-card instead. Many subjects do select the p card, either alone or in conjunction with the q card (A or A and 4 in the example).

Several explanations for this pattern of responses have been postulated. Wason (1966) suggested two possible explanations. Firstly, it is possible that subjects have a "defective" truth table with three rather than two truth values: true, false, and irrelevant. For an implication rule, any case involving p (a vowel) is potentially informative but any case involving $\bar{p}$ (a consonant) is irrelevant. Even though there is evidence for the use of this type of "defective" truth table (Johnson-Laird & Tagart, 1969), this would not have an effect on the relevant cards which have to be chosen to determine the truth of this rule.

The second hypothesis was that subjects were conforming to a verification bias. In other words, subjects tend to confirm rather than disconfirm the rule by choosing the cards that would conform to the rule. This would lead to the choice of the true instances (p and q). Wason had observed a similar verification bias in an inductive reasoning problem (Wason, 1960).

Evans (1972) suggested that a much more primitive process was taking place in this task. He suggested that

subjects were under the influence of a "matching bias" rather than a verification bias, whereby they would merely choose those cards that were mentioned in the rule. To distinguish between these two hypotheses, Evans and Lynch (1973) manipulated the presence of negatives in all possible combinations within the implication rule. This manipulation separated the usefulness of a particular instance as a potential verifier from its matching status as a card named in the rule (Evans, in press). The results supported the matching bias hypothesis over the verification hypothesis, and this result was further supported by Manktelow and Evans (1979).

It is possible that the use of abstract materials in the task inhibits a combinatorial analysis strategy because of the essentially meaningless nature of the problem. Performance on the task could possibly be facilitated by the use of thematic materials. In fact, Wason and Shapiro (1971) observed vastly improved performance when they employed rules employing thematic materials such as "Everytime I go to Manchester I travel by car." In addition, Johnson-Laird, Legrenzi and Sonino Legrenzi (1972) asked subjects to imagine that they were postal workers and to determine if rules such as "If a letter is sealed, then it has a 50 lire stamp on it," were violated. In these studies, performance increased to 62.5% correct and 81% correct, respectively. However, attempts to replicate these results have been less than successful (Brown, Keats, Keats & Seggie, 1980; Griggs

& Cox, in press, Experiments 1 and 2; Manktelow & Evans, 1979, Yachanin, 1980, but see Bracewell & Hidi, 1974, Pollard, 1981; and Van Duyne, 1974).

Both Manktelow and Evans (1979) and Griggs and Cox (in press) hypothesized that thematic materials would facilitate performance only if they served in a memory-cuing capacity. In other words, simply presenting subjects with thematic materials will not lead to correct responding. Only if the materials, the rule, and the falsifying instance are specifically a part of the subject's past experience will correct performance occur. Griggs and Cox (in press, Experiment 3) tested this hypothesis by presenting subjects at the University of Florida with the rule, "If a person is drinking beer, then the person must be over 19 years of age," a law in Florida at the time of the study. Performance increased dramatically to about 70%, but correct performance was not transferred to the abstract selection problem when given immediately afterwards. This suggests that memory cuing is the cause of improved performance rather than the elicitation of logical processing.

Regardless of the processes involved in answering the selection problem, it is obvious that most subjects are not using an appropriate combinatorial truth table analysis to choose the cards to turn over. The common pattern of choosing the p card alone or the p card and the q card together does not conform to the use of an implication rule. This

pattern also does not conform to a biconditional rule, which is a common interpretation of an "if--then" statement (e.g., Moshman, 1979; Rips & Marcus, 1977). The rule would then be interpreted to mean "If there is a vowel on one side of the card, then there is an even number on the other side, and if there is an even number on one side of the card there is a vowel on the other side." If a biconditional interpretation had been given to the rule, all four cards would have to be turned over. Both the $p\bar{q}$ case and the $\bar{p}q$ case will disconfirm a biconditional rule (see Table 1, Rule 10). Thus, the p card (A) would have to be turned over to check for the presence of $\bar{q}$ (7), the $\bar{q}$ card (7) would have to be turned over to check for the presence of p (A), the $\bar{p}$ card (D) would have to be turned over to check for the presence of q (4), and the q card (4) would have to be turned over to check for the $\bar{p}$ (D).

Wason's THOG Problem

Another logical rule that has been investigated within a problem solving framework is exclusive disjunction. This problem was designed by Wason (1977b) and is called the "THOG" problem. As in the selection task subjects must postulate hypotheses, but in the THOG problem they must carry out a combinatorial analysis for each hypothesis to solve the problem.

The task is as follows. The subject is presented with four designs--a black diamond, a white diamond, a black circle, and a white circle. They are then informed that the experimenter has written down one of the shapes (diamond or circle) and one of the colors (black or white). A name (THOG) is related to the designs by the following rule: "If any one of the designs includes either the color I have written down or the shape I have written down, but not both, then it is called a THOG." Subjects are then told that the black diamond is a THOG and that they must decide if each of the remaining shapes is a THOG, is not a THOG, or that there is insufficient information to decide. Performance on this problem is typically slightly better than on the selection problem--from 10%-30% of the subjects gave the correct answer (e.g., Newstead, Griggs & Warner, Note 2; Wason & Brooks, 1979).

Wason (1977b) designed this problem so that if the formation of hypotheses along with a combinatorial analysis is not carried out, a "mirror image" of the correct answer will be given. In fact, this erroneous answer--that the white diamond and black circle are THOGs (or indeterminate) and that the white circle is not a THOG--accounts for over 50% of the errors made and is referred to as the "intuitive" error (Wason & Brooks, 1979). The correct answer to the problem is that the white diamond and the black circle are not THOGs and the white circle is a THOG. The logic of the

problem is as follows. We know that the black diamond is a THOG (given). Thus, because of the stated rule, the experimenter must have written down either black and circle or white and diamond. If the experimenter wrote down black and circle, the white diamond cannot be a THOG because it has neither attribute, the black circle cannot be a THOG because it has both attributes, and the white circle must be a THOG because it has only one attribute. The same conclusion follows in the alternative case. If the experimenter wrote down white and diamond, the white diamond cannot be a THOG because it has both attributes, the black circle cannot be a THOG because it has neither attribute, but the white circle must be a THOG because it has only one attribute. Thus, for both possible combinations the white circle is a THOG and the black circle and white diamond are not THOGs.

There are several hypotheses as to why subjects fail to carry out the analyses necessary to solve the problem. One set of hypotheses focuses on the structure of the task itself. It is possible that a subject's attention is drawn to the instantiated design--the black diamond--and that they reason from this specific instance rather than from what must have been written down based on this instance and the rule. This behavior is suggested by the prevalence of the "intuitive" error and is very similar to the matching bias observed in the selection task. Subjects call any design that is either black or a diamond a THOG and a design with neither of these attributes not a THOG (Griggs & Newstead, Note 3).

An additional problem related to this approach is that subjects may be unable to separate out the attributes of the design and attempt to reason about the design itself. This leads to a conceptual difficulty because the two instances of a THOG necessarily have no property in common (Wason, 1978).

It is also possible that the abstract terms that are utilized in the THOG problem lead to difficulties in logical reasoning. However, merely phrasing the problem in realistic terms does not improve performance (Cordell, 1978; Newstead et al., Note 2, Experiments 1 and 2), but performance is improved by making the correct answer congruent with experience (Newstead et al., Note 2, Experiment 3). There is strong support, however, for the hypothesis that this improvement is not due to improved logic but to memory cuing. When Newstead et al. (Note 2, Experiment 3) asked subjects to justify or explain their answers to the experience-congruent THOG problem, only three of the 13 subjects who answered correctly gave answers that referred to some sort of logic. The other subjects referred specifically to prior experience or knowledge.

In addition, Newstead et al. (Note 2, Experiment 4) found that eight-year-old children demonstrated excellent performance on an experience-congruent THOG task. Since most researchers would agree that these children do not possess the necessary logic structures for a combinatorial

analysis, their behavior must be the result of a nonlogical response bias, namely memory cuing.

Various therapeutic procedures have been utilized as well, with no improvement in performance. Wason and Brooks (1979) had subjects produce the possible hypotheses, but performance still did not improve. Other variations include specifying an example as not a THOG (Griggs & Newstead, Note 3), and separating the attributes from a particular instance (e.g., Cordell, 1978, Newstead et al., Note 2), with little difference in correct performance rate. As with the selection task, it is obvious that most subjects do not appear to be utilizing a truth-table combinatorial analysis in attempting to solve the THOG problem.

CHAPTER 4
GOAL OF DISSERTATION

The major characteristics of formal operations are the ability to formulate and test hypotheses, to isolate and control variables, to carry out combinatorial analyses, and to correctly interpret the results of experimentation. One of the least investigated aspects of formal operations is the hypothesized propositional logic basis of these abilities. The goal of this dissertation is to investigate this hypothesis in two ways: (1) by examining whether performance on traditional Piagetian tasks and on hypothetico-deductive problems which are based directly on propositional logic rules is related and (2) by examining performance on Piagetian tasks employing specific logic probes.

Experiment 1 examines the relationship between performance on paper-and-pencil versions of traditional Piagetian tasks and performance on two propositional logic tasks--the selection task and the THOG problem. Wason (e.g., 1977a) has suggested that the poor performance on these logic tasks, especially the selection task, is enough evidence to seriously undermine the logical bases of Piaget's theory of formal operations. It seems theoretically impossible for propositional logic to underlie formal operational behavior but not a task based more directly on a logical rule, as are both the

relation task and the THOG problem. If this ability underlies performance on both of these types of tasks, performance on the tasks should be correlated.

Three additional factors that have been linked to formal operational ability will be investigated as well. It is possible that these factors will influence performance on one or both of the types of problems being investigated. Past classroom experience will be investigated by recording the number of classes and laboratories in mathematics, chemistry, physics and logic. Scholastic Aptitude Test scores will be recorded as a measure of general ability. In addition, the Group Embedded Figures Test will be administered as a measure of field dependence/independence.

Experiment 2 employs paper-and-pencil versions of two traditional Piagetian tasks, but specific questions based on propositional logic and isolation and control of variables are utilized. Not only does this procedure control for possible ambiguity of task instructions, but it allows the use of reasoning congruent with propositional logic even if this logic would not be expressed spontaneously. Thus, this experiment employs a more sensitive measure of the use of propositional logic than the traditional Piagetian task. As in Experiment 1, the relationship of the number of relevant classes and laboratories taken and Scholastic Aptitude Test scores will also be investigated.

CHAPTER 5
EXPERIMENT 1

Experiment 1 was conducted to investigate the relationship between formal operational ability, logical ability, field dependence/independence, past classroom experience, and a measure of general overall ability. Many researchers have hypothesized relationships between various combinations of these factors (e.g., Flexer & Roberge, 1980; Neimark, 1975; Piaget, 1972; Pulos & Linn, in press; Wason, 1977), but as of yet no within-subjects experiment has investigated all of them. Specifically, the examination of the relationship between performance on the formal operational tasks and on the logic tasks is intended as a test of Piaget's assertion that propositional logic underlies formal operational abilities (Inhelder & Piaget, 1958).

Although the formal operational tasks used in this experiment were group-administered paper-and-pencil tests, they were designed to be as close as possible to those used in Inhelder and Piaget's (1958) experiments. Recently, some researchers have been focusing on only one aspect of formal operational thinking (e.g., Barratt, 1975; Kuhn & Brannock, 1977), but the tasks in the present experiment allowed for a wider range of abilities to be expressed. In addition, permutation ability was measured separately.

The logic tasks administered were the abstract selection problem and THOG problem. These problems indirectly test propositional logic abilities in a problem-solving framework, as Piaget has proposed for the solution to his formal operational tasks. One question of particular interest is whether subjects who do well on these tasks also do well on the formal operational tasks. This result would support Piaget's hypothesis that propositional logic underlies formal operational abilities.

The Group Embedded Figures Test (Oltman, Raskin & Witkin, 1971) was administered as a test of field dependence/independence. There is evidence that this trait correlates with formal operational ability (e.g., Neimark, 1975). It is also plausible that the ability to extract important variables from irrelevant background "noise" (i.e., field independence) is important for successful performance on the logic tasks.

In addition, past classes and laboratories taken in mathematics, chemistry, physics, and logic in both high school and college were examined. As Piaget has stated (Piaget, 1972), specific past experience may be the most important indicator of performance within a particular domain Since the problems utilized in this experiment are science-based, the above classes and laboratories seemed to be the most appropriate ones to investigate.

Finally, a measure of general abilities--Scholastic Aptitude Test (Verbal) [SAT(V)] and Scholastic Aptitude Test (Quantitative) [SAT(Q)] scores--was recorded. It is possible that a student will perform well on all tasks because of some more general ability rather than because of logical ability or scientific reasoning.

Method

Subjects

Ninety-one students at the University of Florida served as subjects in this experiment as partial fulfillment for course requirements in introductory psychology. Forty-six subjects were female, and 45 were male.

Materials

As a measure of field independence, the Group Embedded Figures Test (GEFT) developed by Witkin and his colleagues (Oltman, et al., 1971) was administered to all subjects. To assess formal logic abilities, two problems were formulated. One problem, a version of Wason's selection task, is based on the propositional logic rule of implication. In addition, subjects were asked to justify their answer after it had been given. This problem along with the justification sheet is given in Appendix A. The second problem, a version of Wason's THOG problem, is based on the propositional logic

tule of exclusive disjunction. Subjects were also asked to justify their answer to this problem. The justification sheet and the problem are given in Appendix B.

Four paper-and-pencil problems were developed as a measure of formal operations. Three problems were based very closely on three of the 15 problems described in Inhelder and Piaget (1958)--the bending rods problem, the pendulum problem, and the combination of colorless chemicals problem. The format of the problems was based on a group-administered, formal-operational task developed by Faust and Golembowski (Note 4). Subjects are presented with a verbal and pictoral representation of the problem and are asked to describe in writing as completely and exactly as possible how they would solve the problem.

These three problems were presented together in a booklet preceded by an information and instruction page. On this page, subjects were asked to list all the classes and laboratories in logic, mathematics, chemistry and physics that they had taken in high school and in college. The instructions informed the subjects of the nature of the task and the importance of answering the questions completely and exactly. This problem booklet is given in Appendix C.

In addition, a permutation problem based on Neimark's (1975) license-plate problem was administered. In this problem, subjects are asked to produce all possible license plate numbers from a given four digit license plate. Then,

subjects are asked to try and discover a general rule that will predict the number of possibilities for any given number of digits This problem is given in Appendix D.

Procedure

Subjects were run in same-sex groups of from eight to 10 subjects each. The experiment consisted of two one-hour sessions separated by one day. In one session, subjects completed the GEFT, the selection problem, and the THOG problem. The other session consisted of the booklet of formal operations tasks and the permutation task. A pilot study had indicated that most subjects would complete each task group in approximately an hour.

The order of the tasks was counterbalanced both over sessions and within sessions by sex. The order of the Piagetian tasks within the test booklet was randomized as well. The general layout of the experiment is given in Table 2.

All tasks were self-paced except for the GEFT, which is a timed test. The subjects were reminded by the experimenter to answer the questions as exactly and completely as possible before each session. They were also reminded that the appropriate tasks were self-paced and that they were allowed as much time as necessary to complete the task.

Table 2.

The Order of Tasks in Experiment 1

Group	N	Session 1	Session 2
Females			
1	9	Piagetian tasks Permutation	GEFT THOG Selection
2	10	Selection THOG GEFT	Permutation Piagetian tasks
3	9	Permutation Piagetian tasks	Selection THOG GEFT
4	8	GEFT THOG Selection	Piagetian tasks Permutation
5	10	Piagetian tasks Permutation	GEFT THOG Selection
Males			
1	8	THOG Selection GEFT	Permutation Piagetian tasks
2	9	Permutation Piagetian tasks	Selection THOG GEFT
3	9	GEFT THOG Selection	Piagetian tasks Permutation
4	9	Piagetian tasks Permutation	GEFT THOG Selection
5	10	Selection THOG GEFT	Permutation Piagetian tasks

Results

First, the descriptive statistics for this group of subjects will be discussed. The mean age was 19.27 years (SD=1.41). The mean score on the GEFT was 12.16 points (SD=4.95). The highest actual score was 18 points and the lowest actual score was 0 points. Only 84 of the 91 subjects had taken the SAT exams. The mean for the SAT (Q) test was 507.62 points (SD=108.56). The highest actual score was 710 points and the lowest actual score was 270 points. The mean score for the SAT (V) test was 481.67 points (SD=105.99). The highest actual score was 720 points and the lowest actual score was 270 points. The mean number and range of classes taken in each subject area is presented in Table 3. As can be seen in Table 3, the mean number of all classes taken was below one class, except for the subject of mathematics

Piagetian Problems

Permutation Problem. The permutation problem was scored as correct if the subjects successfully permuted the four numbers given in the license plate to produce the 23 additional numbers. Seventy-one of the 91 subjects (78%) did so. The other subjects only partially permuted the numbers. Subjects were also asked to deduce (or produce) the general formula for permutations. Twenty-six of the subjects (29%) were able to do this. The remaining subjects either

Table 3.

Mean Number, Standard Deviation, and Range of Classes Taken by Subjects in Experiment 1

Class	Mean	Standard Deviation	Range
Logic (HS)[a]	.02	.15	0-1
Logic (C)[b]	.13	.34	0-1
Mathematics (HS)	3.48	1.17	0-6
Mathematics (C)	2.04	1.64	0-9
Mathematics lab (C)	.04	.21	0-1
Chemistry (HS)	.97	.62	0-3
Chemistry lab (HS)	.72	.66	0-2
Chemistry (C)	.91	1.65	0-8
Chemistry lab (C)	.70	1.43	0-7
Physics (HS)	.52	.58	0-3
Physics lab (HS)	.34	.56	0-3
Physics (C)	.43	.99	0-4
Physics lab (C)	.40	.95	0-4

[a]HS refers to high school

[b]C refers to college

gave the wrong formula, stated that they did not know the answer, or did not answer the question.

Traditional Piagetian Problems These problems were scored according to the scoring system given in Table 4. This scoring system is based on performance descriptions given by Inhelder and Piaget (1958) and the performance criteria of Martorano (1977) and Peel (1971). The problem booklets were taken apart, and the data for each problem were scored independently of each other. Two people scored all of the answers, and interrater reliability was .90. A mutual decision was reached on all problems on which the raters disagreed, bringing interrater reliability to 1.0.

The means, standard deviations, and number of subjects performing at each level are given in Table 5. These data will be compared with those of Martorano (1977) because of the similarity of the tasks and scoring system. Martorano's data for twelfth grade subjects are also given in Table 5 (adjusted down for a scale difference of one point). The means for the present study are somewhat lower than those reported by Martorano. This difference in overall level of difficulty could be due to several factors. For example, the subject pools differ in both number and makeup--Martorano's twelfth grade subjects were 20 females. Forty-five males as well as 46 females participated in this study. Thus, it is possible that this larger number of subjects resulted in a more representative sample of the population.

Table 4.

The Scoring System for the Three Traditional Piagetian tasks in Experiment 1

Level of Behavior	Category	Characteristics of Answers[a]
Preoperational	0	No answer; Random trial and error tests; No logical process made; No concept of variable isolation, Refusal to accept the task (i.e., tautological answers, irrelevant answers, reasoning on arbitrary or irrelevant grounds)
Early Concrete	1	Some concept of isolation of variables (e.g., elimination of one variable but failure to unconfound the other possible factors, or drawing conclusions on incompletely isolated or combined variables), no mention of control of variables
Late Concrete	2	No mention of control of variables; May carry out sufficient binary combinations but be unable to make full use of the information obtained; Unsystematic combinations
Early Formal	3	Systematically tests out dimensions of variation occurring in the task; Isolates operative variable and logically excludes inoperative variables, Answer is logically consistent
Late Formal	4	In addition to testing of hypotheses, the subject generalizes the results and forms a rule that applies to all similar situations, Makes predictions and/or possible conclusions from all possible outcomes.

[a]This scoring system is based on Martorano (1977) and Peel (1971), and the performance criteria of Inhelder and Piaget (1958).

Table 5

The Number of Subjects at Each Level and Mean Scores and Standard Deviations for the Three Traditional Piagetian Problems Given in Experiment 1

Problem	Level 0	Level 1	Level 2	Level 3	Level 4	Present Study[a] Mean Score	(SD)	Martorano (1977)[b] Mean Score	(SD)
Bending Rods	16	15	35	23	2	1.78	(1.08)	2.70	(.80)
Pendulum	21	5	15	46	4	2.08	(1.29)	2.60	(.60)
Combination of Colorless Chemicals	20	8	3	60	0	2.13	(1.27)	3.00	(.56)

[a]N = 91 subjects

[b]N = 20 subjects. These data have been adjusted down one point to compensate for scale differences, and are only for the 12th grade subjects.

Further evidence for this hypothesis comes from the standard deviations for each mean--Maratorano's subjects performed with standard deviations ranging from .56 to .80 on these tasks. The subjects in this experiment performed with standard deviations ranging from 1.08 to 1.29, reflecting a much wider range in performance.

However, the most obvious difference between these two studies is in the form of the problems. The problems in this study were paper-and-pencil tasks. Martorano's subjects used actual task apparatus, and both verbal and nonverbal responses were scored. It is possible that even though subjects were given as much time as necessary to answer each problem, a totally written response is more influenced not only by level of cognitive ability, but also by motivation, verbal ability, and so on, than an active physical response. It is also possible that behaviors which would naturally occur in making physical responses are not thought to be necessary in a written response. For example, subjects may not feel that it is necessary to state that a pendulum should be held at the same level before letting go during each test of a particular factor, and yet would do so during an actual experiment. Scorers attempted to control for this type of error by giving subjects credit if the concept of control was mentioned for at least one factor that should be controlled, but there are many possible potential aspects of relevant behaviors that subjects may not have

written down that are not as easily checked for. Analyses investigating the relationship between performance on these tasks and verbal ability will be discussed later in the section called Analyses Investigating the Relationship Between Factors.

One related trend in the data for the bending rods task, however, was easily discernable. Many subjects stated that they would test all possible combinations of rods, but they did not state which factors they would be investigating with each combination. This error placed these subjects at Level 2, in which a characteristic response is to carry out sufficient binary combinations but be unable to make full use of the information obtained. It is impossible to tell whether or not subjects actually did not know how to make sufficient use of the information obtained or simply did not write down the answer. This was the only discernable trend of this type in the data.

Even though the mean scores are concrete operational or less, an examination of the pattern of responses for each task reveals that most subjects scored in at least the early formal operations level for both the pendulum and the combination of colorless chemicals task. The number of subjects scoring in Level 2 for the bending rods task was the result of the type of error described above. Thus, even though mean scores were lower than those reported previously by Martorano (1977), the number of subjects scoring in the

formal operational levels is relatively high 25 of 91 subjects in the bending rods task (27%), 50 of 91 subjects in the pendulum task (55%), and 60 of 91 subjects in the chemicals task (66%). These percentages, along with the percentage correct for permutations, are similar to those reported by Martorano (1977) except for the bending rods problem. Sixty percent of the twelfth grade subjects in Martorano (1977) performed at the formal operational level.

Propositional Logic Problems

The data for each propositional logic task were scored independently so that performance on one task would not influence the scoring of the other task. Answers were scored as correct (2), partially correct (1), or incorrect (0). For both tasks, any response pattern which was partially correct with a correct justification for the correct part of the answer was scored as partially correct. For the selection task, answers were scored as correct if both the p and $\bar{q}$ cards were chosen, partially correct if only the p card was chosen (no one chose $\bar{q}$ only), and incorrect if any other answer was given. For the THOG task, answers were scored as correct if the response pattern BBA was given, partially correct if the response patterns CCA, BBB or BBC were given, and incorrect if any other answer was given. These partially correct patterns were the only ones which occurred. (See Appendix B for an explanation of the response choices.)

Subjects were also asked for a justification for each answer given. All justifications were scored as correct (1) or incorrect (0) based on their adherence to the appropriate logical rule. The justification for every card had to be correct for a justification score of one to be given. If subjects changed their answer, the new answer was recorded as their final answer. Thus, if while justifying an incorrect answer the subject realized his or her mistake and rectified it with a correct answer and justification, full credit was given for the answer.

Selection Task. The data for the selection problem are given in Table 6. These results are like those typically found in previous studies using the abstract selection problem (e.g., see Griggs & Cox, in press). The number of subjects choosing the correct answer is inflated, however, because of the number of subjects changing their answer to the correct one. Only five of 91 subjects originally chose the correct cards, which is closer to the typical results.

Many subjects chose the p and q cards, which is usually the most prevalent answer, presumably because of a verification or matching bias (e.g., Evans & Lynch, 1973; Wason, 1966). The second most common answer was to turn over all of the cards, and the correct answer (p and $\bar{q}$) was the third most prevalent answer. Subjects justified turning over all four cards in one of two ways. Seven subjects stated that there was no way of knowing what was on the other side of

Table 6.

The Response Patterns for the Selection Task in Experiment 1

Cards Chosen	N
p, q	24
p, $\bar{p}$, q, $\bar{q}$	16
p, $\bar{q}$[a]	15
p[b]	9
$\bar{p}$, $\bar{q}$	9
$\bar{p}$, q	4
p, $\bar{p}$	1
none	4
Total	91

[a]Correct

[b]Scored as partially correct.

any card unless it was turned over. The remaining nine subjects mentioned looking for a specific symbol on the back of each card, which can be more readily identified as the use of a biconditional rule. As discussed in Chapter 2, a common interpretation of an implication statement is in terms of a biconditional rule. For a biconditional rule, all four cards must be turned over in order to check for falsifying cases. It is also possible that some or all of the seven subjects mentioned above may have been using a biconditional rule as well.

All subjects who made the correct selection justified their answer in accordance with the implication rule. The justifications for the remaining subjects' choices were not identifiable as conforming to any propositional logic rule, except for those who appeared to be using a biconditional rule.

Since subjects in this experiment were required to justify their answers and were instructed to change their answer if they felt it necessary to do so at that point in the experiment, the number of subjects who did change their answer was examined. Twenty-seven subjects changed their answer to another wrong answer, ten subjects changed their answer to the right answer, and one subject changed her answer from the right to the wrong answer. Thus, a total of 38 of 91 subjects (42.75%) changed their answer. In fact, of the 15 subjects who answered correctly, ten did so after

beginning the justification process. The act of justification itself seems to provide insight into the correct answer at least for some subjects.

THOG Task. The response patterns for the THOG problem are presented in Table 7. Again, the responses are similar to data in previous investigations of this problem (e.g., Wason & Brooks, 1979). The majority of incorrect responses correspond to the "intuitive" error--29 of the 91 subjects responded AAB and 13 responded CCB for a total of 46.0% of the subjects reasoning in this manner. Interestingly, the justification for an intuitive error requires the subject to state that a white diamond is a THOG (or indeterminate) because it has one of the characteristics of the THOG (the black diamond), a black circle is a THOG (or indeterminate) because it also has one of the characteristics of the THOG, and the white circle is not a THOG because it has neither of the characteristics. All of the subjects making this error gave this justification, but only one subject mentioned the contradiction in this answer; namely that the original THOG (the black diamond) cannot be a THOG using this reasoning because it has both characteristics. It is not known if the other subjects using the typical intuitive error justification did not notice the contradiction or chose to ignore it.

The 18 subjects who answered correctly justified their answer in terms of the correct reasoning for the exclusive disjunctive rule. Thirteen subjects answered partially

Table 7.

The Response Patterns for the THOG Problem in Experiment 1

Answer	N
AAB	29
CCB	13
BBA[a]	18
CCA[b]	8
CCC	5
AAA	5
BBB[b]	4
other	9
Total	91

[a]Correct

[b]Scored as partially correct. A response pattern of BBC was also scored as partially correct, but only one subject made this response. This subject's answer is classified under "other."

correctly (one subject answered BMC and is included in the "other" category) because they failed to carry out totally the required combinatorial analysis. The remaining 60 answers were not identifiable as conforming to any propositional logic rule.

As for the selection task, the modification of answers during justification was investigated. Ten subjects changed their answer to another wrong answer, 12 changed from a wrong to a right answer, and seven changed from a wrong to a partially right answer, for a total of 29 of 91 subjects, or 31.9%. Twelve of the 18 subjects who answered correctly did so during the justification process. As in the selection problem, at least some subjects experience insight into the problem solution during justification.

Analyses Investigating the Relationship Between Factors

The major purpose of this study was to investigate the relationship between the measured factors. In order to do this, a correlation matrix was formed for the 27 factors using Pearson product-moment correlations. This correlation matrix is given in Table 8. As can be seen in this table, many intercorrelations occur between the variables, which renders interpretation of these correlations impossible because of confounding. For example, sex is significantly correlated with performance on the THOG problem, but the number of classes taken in chemistry during high

Table 8.

The Correlation Matrix for the Factors in Experiment 1

	AGE	SEX	LHS	LC	MHS	MC	MLAB	CHS
AGE	1.00	-.12	-.03	.22*	-.26*	.47*	.15	-.05
SEX		1.00	.15	-.13	-.12	-.20	-.002	-.23*
LHS			1.00	-.06	-.13	-.10	-.03	-.23
LC				1.00	-.02	.21*	-.08	.02
MHS					1.00	.18	-.23*	.36*
MC						1.00	-.14	.21*
MLAB							1.00	.01
CHS								1.00
CLAB								
CCOL								
COLAB								
PHS								
PLAB								
PCOL								
PCLB								
SATQ								
SATV								
GEPT								
SEL								
SELR								
THOG								
THOGR								
PERM								
FORM								
ROD								
PEND								
CHEM								

*$p < .05$

Legend: AGE = age of subject, SEX = sex of subject, LHS = number of logic classes taken in high school, LC = number of logic classes taken in college, MHS = number of mathematics classes taken in high school, MC = number of mathematics classes taken in college, CHS = number of chemistry classes taken in high school, CLAB = number of chemistry laboratories taken in high school, CCOL = number of chemistry classes taken in college, COLAB = number of chemistry laboratories taken in college, PHS = number of physics classes taken in high school, PLAB = number of physics laboratories taken in high school, PCOL = number of physics classes taken in college, PCLB =

	CLAB	CCOL	COLAB	PHS	PLAB	PCOL	PCLB	SATQ
AGE	.04	.29*	[illegible]	.12	.10	.49*	.50*	.12
SEX	-.22*	-.37*	-.30*	-.26*	-.26*	-.28*	.26*	-.31*
LHS	-.17	-.08	-.07	-.13	-.09	-.07	-.06	-.21*
LC	.17	-.14	-.12	.16	.17	-.005	.01	.19
MHS	.44*	.20	.17	.35*	.22*	.001	.01	.38*
MC	.34*	.38*	.42*	.39*	.29*	.56*	.56*	.38*
MLAB	-.16	-.09	-.11	-.19	-.13	-.09	-.09	-.25*
CHS	.62*	.45*	.40*	.44*	.41*	.28*	.28*	.32*
CLAB	1.00	.43*	.40*	.54*	.58*	.28*	.27*	.38*
CCOL		1.00	.96*	.29*	.32*	.61*	.60*	.30*
COLAB			1.00	.28*	.31*	.66*	.65*	.33*
PHS				1.00	.78*	.40*	.41*	.33*
PLAB					1.00	.43*	.43*	.26*
PCOL						1.00	.98*	.21*
PCLB							1.00	.25*
SATQ								1.00
SATV								
GEFT								
SEL								
SELR								
THOG								
THOGR								
PERM								
FORM								
ROD								
PEND								
CHEM								

umber of physics laboratories taken in college. SATQ = score on the SAT (Q) examination, SATV = score on the SAT (V) examination, GEFT = score on the Group Embedded Figures Test, SEL = score on the selection task, SELR = score for reasoning on the selection task, THOG = score on the THOG task, THOGR = score for reasoning on the THOG task, PERM = score on the permutation task, FORM = score for the formula for permutations, ROD = score on the bending rods task, PEND = score on the pendulum task, CHEM = score on the combinations of colorless chemicals task.

Table 8--continued

	SATV	GEFT	SEL	SELR	THOG	THOGR	PERM	FORM
AGE	.23*	.07	.24*	.25*	.10	.09	-.08	-.002
SEX	-.15	-.24*	.02	.02	-.23*	-.15	-.21*	-.01
LHS	-.18	-.14	-.09	-.07	-.09	-.07	-.28*	-.09
LC	.20	.10	.03	.002	.12	.15	-.03	.04
MHS	.14	.11	.08	.07	.11	.19	.15	.18
MC	.31*	.23*	.20	.12	.03	.07	.13	.25*
MLAB	-.11	-.03	.02	.05	.13	.03	-.15	-.02
CHS	.18	.15	.03	-.02	.30*	.21*	.06	.03
CLAB	.36*	.24*	.02	-.04	.17	.21*	.11	.12
CCOL	.17	.10	.14	.10	.17	.13	.02	.03
COLAB	.20	.14	.16	.11	.14	.12	.001	.06
PHS	.36*	.17	.16	.11	.05	.06	.15	.11
PLAB	.34*	.13	.19	.10	.08	.06	.09	.09
PCOL	.22*	.22*	.25*	.14	-.01	-.04	.02	.07
PCLB	.23*	.23*	.26*	.16	-.003	-.02	.03	.07
SATQ	.68*	.52*	.34*	.25*	.39*	.43*	.34*	.34*
SATV	1.00	.55*	.34*	.25*	.30*	.34*	.11	.29*
GEFT		1.00	.23*	.11	.21*	.25*	.08	.17
SEL			1.00	.91*	.32*	.28*	.13	.15
SELR				1.00	.27*	.24*	.09	.11
THOG					1.00	.89*	.04	.03
THOGR						1.00	.12	.07
PERM							1.00	.16
FORM								1.00
ROD								
PEND								
CHEM								

Table 8--continued

	ROD	FEND	CHEM
AGE	-.04	.07	.07
SEX	-.06	-.16	-.17
LHS	-.04	-.13	-.07
LC	-.01	.10	.04
MHS	.18	.11	.14
MC	.14	.23*	.24*
MLAB	-.21*	.07	-.06
CHS	.15	.17	.13
CLAB	.06	.15	.17
CCOL	.14	.17	.17
COLAB	.17	.20	.19
FHS	.13	.20	.24*
FLAB	.11	.19	.25*
FCOL	.19	.18	.28*
FCLB	.18	.17	.29*
SATQ	.33*	.43*	.53*
SATV	.32*	.52*	.52*
GEFT	.21*	.47*	.50*
SEL	.04	.19	.24*
SELR	.01	.16	.16
THOG	.06	.15	.25*
THOGR	.12	.21*	.26*
PERM	.14	.03	.10
FORM	.17	.11	.11
ROD	1.00	.52*	.44*
FEND		1.00	.48*
CHEM			1.00

school is significantly correlated to both sex and performance on the THOG problem. Thus, it is impossible to discern if the effect of sex on the THOG problem is actually an effect due to sex or to the number of chemistry classes taken in high school. Therefore, the significance of partial correlations between the major variables was investigated. Partial correlations allow for a measure of the correlation between two variables, controlling for the effects of all other variables in the model.

Significance tests for partial correlations for the factors of SAT (Q), SAT (V), the selection task, reasoning for the selection task, the THOG task, reasoning for the THOG task, and the four formal-operational tasks were carried out. These tests tested for the significance of the partial correlations between the selected factors and all other factors by testing H_0: $B_1 = B_2 = . . . B_n = 0$ by the Fs for the Type IV sum of squares in a regression equation (Kleinbaum & Kupper, 1978; Helwig & Council, 1979). The significant partial correlations are given in Table 9.

All partial correlations with a significance level of $p<.05$ are reported in Table 9, but because of the number of comparisons that were made a more reasonable level of significance is $p<.01$ to guard against a Type II error. SAT (Q) and SAT (V) scores are highly correlated, as are SAT (Q) and the number of physics classes taken in college. Since the SAT examinations were taken before the physics classes,

Table 9.

Significant Partial Correlations between Selected Factors in Experiment 1

Partial Correlation Factors	Significance Level
Selection task, Reasoning for selection task	0001
THOG task, Reasoning for THOG task	.0001
Bending rods task, Pendulum task	.0006
SAT (Q), SAT (V),	.0008
SAT (Q), Number of physics classes taken in college	.0092
SAT (Q), Number of physics laboratories taken in college	.0240
THOG task, Number of chemistry classes taken in high school	.0257
THOG task, Sex	.0281
Pendulum task, Number of mathematics laboratories taken in college	.0388
Reasoning for selection task, Age	.0401
SAT (Q), Chemical combinations task	.0423
Selection task, Sex	.0439

it seems reasonable to conclude that subjects who score high on the SAT (Q) exam tend to take more physics classes in college than those students who score lower on the exam.

The reasoning for both the propositional logic tasks and performance on these tasks are highly correlated as well. All subjects who answered correctly on both tasks justified their answers correctly. However, the relationship between the two tasks themselves was not significant. Examination of the subjects' performance on both tasks reveals that six of the 91 subjects answered both tasks correctly, 52 of the subjects answered both incorrectly, eight answered the THOG problem correctly but missed the selection problem, six answered the selection problem correctly but missed the THOG problem, and the remaining 18 subjects answered one, or both of the two problems partially correctly in various combinations.

Two of the formal operations tasks were highly correlated, the bending rods task and the pendulum task. Similar results have been reported before (e.g., Danner & Day, 1977; Martorano, 1977). However, no significant correlation was observed between any of the other formal operational tasks. This may be because the different tasks measure different aspects of formal operational ability. Inhelder and Piaget (1958) describe the bending rods task and the pendulum task as measuring multiplicative compensation, and the high degree of correlation between performance on these two tasks

may reflect this common necessary ability. In addition, the chemicals task measures only combinatorial ability, while the license plate task measures permutational ability. The bending rods task and the pendulum task require more than combinations or permutations to be scored at a high level (as demonstrated by the Level 2 scores on the bending rods task mentioned previously)

One of the more striking aspects of the data in Table 5 is the lack of correlations between the logic tasks and the Piagetian tasks. In addition, the correlation between these tasks and past classes is marginal or nonexistent, as is the relationship between any of the tasks and SAT (Q), SAT (V), or GEFT scores. The lack of correlation between the logic tasks and any of the Piagetian tasks may indicate the existence of separate underlying cognitive abilities. Since the logic tasks are based directly on a specific propositional logic rule but the formal operational tasks are only theoretically linked to propositional logic, alternative abilities (such as isolation and control of variables) appear to be more crucial for successful behavior on traditional formal operational tasks. Lawson, Karplus, and Adi (1978) found similar results using a principle-components analysis on their data. Lawson et al. (1978) administered seven formal operational tasks, two of which were propositional logic tasks, including a version of Wason's selection task. The five tasks used to measure three formal

abilities (proportions, correlations and probability) had principle loadings on one factor and the two propositional logic tasks had principle loadings on a second factor. Lawson et al. also argue that the isolation and control of variables may be the principle ability that constitutes formal thought. They further suggest that the term "formal thought" be replaced by the term hypothetico-deductive thought since formal logic does not appear to underlie thought in this stage.

The lack of relationship between science training and formal operational ability has been discussed by several researchers, as mentioned in Chapter 1. Even though Piaget (1972) modified his theory to accommodate the effects of past experience, apparently the experience must be more recent or an overlap between the person's everyday experience and the framework in the problem must occur. Recent studies based on this hypothesis, however, have not always been successful (e.g., Capon & Kuhn, 1979). Also, only the number of classes taken was measured in this experiment, and not the grade received or the understanding obtained in each course, which probably is the critical variable.

Previous studies have indicated a relationship between field dependence/independence (as measured by the GEFT in this experiment) and formal operational abilities (e.g., Linn, 1978; Neimark, 1975, Pascual-Leone, 1970). Other studies have not found a consistent relationship between these

two factors (e.g., Shaini, 1973). Flexer and Roberge (1980) demonstrated that the relationships between formal operational abilities and field dependence/independence were due primarily to their common overlap with IQ (as measured by Lorge-Thorndike Intelligence Tests). Although the SAT examinations are not measures of IQ, they are measures of general ability, so it was unusual to find no relationship between at least some of these factors. Therefore, an additional analysis of the data was performed.

As a final measure of the relationships between the factors measured in this study, a series of stepwise regressions were carried out. This analysis tests for the best predictor(s) of the variable of interest from a set of n variables. All possible models are tested for the best predictor or predictors of performance, i.e., those variables that account for the most variance as measured by r^2.

Because of the intercorrelations between many of the factors, measures were collapsed into their respective overall factor type. The SAT (Q) and SAT (V) measures were combined to form a measure called SAT. The numbers of classes and laboratories taken in all four areas were collapsed into a total measure of classes taken. The permutation and permutation-formula scores were combined, as were the selection task and selection-reasoning scores and the THOG task and THOG-reasoning scores. The measures of sex, GEFT, and the scores on the bending rods, pendulum, and combination of

colorless chemicals tasks were not collapsed. This left a total of 10 factors. Stepwise regressions were run for the dependent variables of selection score, THOG score, permutation score, bending rods score, pendulum score and combination of colorless chemicals score to discover the best predictors for performance on these tasks.

The best one-, two-, and three-variable models for these factors are given in Table 10. No variable entered the remaining models at a significant level. As can be seen in Table 10, SAT is the best predictor of performance on both of the logic tasks and two of the formal operations tasks (permutations and chemical combinations).

For each of the logic tasks, the other logic task enters into the model second, and sex enters third. However, only approximately 20% of the variance is accounted for by the best three predictors of performance on the logic tasks. None of the remaining variables were significant predictors of performance on these two tasks.

Interestingly, the logic tasks are better predictors for permutational ability than the traditional Piagetian tasks, though they are not significant predictors. SAT is the only significant predictor of permutational ability, and only 18% of the variance is accounted for by the best three-variable model.

The best predictors for each of the three traditional formal operational tasks are SAT, one of the other three

Table 10.

The Best One-, Two-, and Three-Variable Models for Selected Factors in Experiment 1

Factor	One-Variable Model (r^2)	Two-Variable Model (r^2)	Three-Variable Model (r^2)
Selection task	SAT[a,b] (.12)	SAT[a] THOG Task (.16)	SAT[a] THOG task[a] Sex (.20)
THOG task	SAT[a] (.15)	SAT[a] Selection Task (.19)	SAT[a] Selection Task[a] Sex (.22)
Permutations	SAT[a] (.15)	SAT[a] Selection Task (.17)	SAT[a] Selection Task THOG task (.18)
Bending rods	Pendulum[a] (.29)	Pendulum[a] Chemical Comb. (.32)	Pendulum[a] Chemical Comb.[a] GEFT (.33)
Pendulum	Bending rods[a] (.29)	Bending rods[a] SAT[a] (.42)	Bending rods[a] SAT[a] GEFT (.45)
Chemical Combinations	SAT[a] (.34)	SAT[a] Pendulum[a] (.40)	SAT[a] Pendulum[a] GEFT (.43)

[a]Significant at $p<.05$

[b]SAT is a combined score of SAT (Q) and SAT (V)

tasks, and GEFT score (which never accounts for a significant amount of variance). From 33% to 45% of the variance is accounted for by the best three-variable model for the tasks, and no other variables were significant predictors of performance.

The difference in amount of variance accounted for by the best three-variable models for the logic problems versus the Piagetian problems could be because of two factors. One is that subjects performed at a much lower level on the logic tasks than on the Piagetian tasks; and since subjects performed uniformly badly, none of the factors predict performance very well. On the other hand, it is possible that unknown variables not entered into the analysis contribute more to performance on the logic tasks; and variables in the analysis were sufficient to contribute significantly to the Piagetian tasks. This is support for the hypothesis that different abilities underlie the two types of tasks, as proposed by Lawson, Karplus, and Adi (1978). Further support is shown by the particular variables entering into the models--although SAT is a predictor for both types of tasks, similar tasks are better predictors for each other than dissimilar tasks.

Since SAT was a better predictor than the total number of classes taken, general ability appears to be more important for prediction of performance on these tasks. As mentioned before, however, different results may have occurred

if understanding of the subject matter had been measured rather than merely the number of classes taken.

If two different types of cognitive abilities underlie the logic tasks and the traditional Piagetian tasks, then typical performance on the selection task, for example, is not as damaging to Piaget's theory of formal operations as Wason (1977a) claims. However, it does appear that the use of propositional logic does not underlie performance on traditional Piagetian tasks, because the logic tasks are based strictly on propositional logic rules and poor performance on these tasks does not preclude good performance on the Piagetian tasks. The isolation and control of variables appears to be the most important factor for successful performance on Piagetian tasks using traditional performance criteria. Also, the lack of correlation between performance on the selection task and on the THOG task is evidence that the two rules on which they are based (implication and exclusive disjunction) are not linked together psychologically by the INRC group.

The importance of SAT scores for predicting behavior on the logic tasks and on the Piagetian tasks does provide a link between them, but a more important factor seems to separate task performance. It remains to be seen whether this factor is task differences or truly a cognitive ability difference, such as amount of working memory necessary to solve the task, heuristic choice, or some other ability.

CHAPTER 6
EXPERIMENT 2

Experiment 1 demonstrated that the use of propositional logic is not necessary for successful performance on traditional Piagetian tasks. Experiment 2 was conducted to probe more specifically for the use of propositional logic in solving such problems as well as for the use of more general formal operational abilities. Two problems from Experiment 1--the bending rods problem and the pendulum problem--were modified for use in this experiment. These particular problems were chosen because they were suited for the modifications introduced in this experiment and because performance on these tasks was highly related in Experiment 1. Formal operational subjects should be able to answer specific logical probes even if they do not spontaneously produce evidence for the use of propositional logic in solving problems like those in Experiment 1.

The difference in performance on propositional logic tasks and on the Piagetian tasks in Experiment 1 may have been due to task differences. Although evidence does not exist for the use of propositional logic in solving open-ended traditional Piagetian problems, it is possible that subjects will use propositional logic if the specificity of the questions makes it necessary for successful performance. The

probes used in Experiment 2 asked for specific knowledge of isolation and confounding of variables, presented results for evaluation, and gave specific hypotheses to be tested Thus, these probes permitted an examination of the use of propositional logic in solving traditional formal operational problems.

Method

Subjects

Forty-six students at the University of Florida served as subjects in this experiment as partial fulfillment for course requirements in introductory psychology classes. None of the subjects had participated in Experiment 1. Twenty-four of these subjects were female, and 22 were male.

Materials

The bending rods problem and the pendulum problem administered in Experiment 1 were modified for this experiment. Ten questions were formulated for the bending rods problem and nine questions were formulated for the pendulum problem. Each problem was designed to take approximately an hour to complete. Both of these problems are given in Appendix E.

The questions in each problem were formulated to assess both general formal operational abilities as well as the ability to use specific propositional logic rules. As can

be seen in Appendix E, the first five questions for each problem are very similar. These probes are based on some questions from a group-administered test of formal operations developed by Tomlinson-Keasey and Campbell (Note 5). The first question tests spontaneous knowledge of the relevant factors for each problem. Questions 2 and 3 test for the ability to isolate and test variables. Questions 4 and 5 test for the ability to recognize the confounding of variables.

The remaining questions were designed to tap the use or nonuse of propositional logic rules. In the bending rods problem, the hypotheses to be tested and the results are stated such that in logical nomenclature p and q could refer to the two factors of length and material (see Dawes, 1972 and Ginsburg & Opper, 1979). For example, if p refers to long, $\bar{p}$ refers to short, and if q refers to brass, $\bar{q}$ refers to steel. Thus, the four possible combinations of these factors result in the four primary conditions in the binary propositional logic table:

pq (the long brass bar)

$p\bar{q}$ (the long steel bar)

$\bar{p}q$ (the short brass bar)

$\bar{p}\bar{q}$ (the short steel bar).

The results of experimentation are then represented by a 1 (for a positive result, such as the bar bends), or a 0 (for a negative result, such as the bar does not bend). Each

pattern of results corresponds to a propositional rule in Table 1. For example, suppose the following results occurred:

pq 1 (the long brass bar bends)

$p\bar{q}$ 1 (the long steel bar bends)

$\bar{p}q$ 0 (the short brass bar does not bend)

$\bar{p}\bar{q}$ 0 (the short steel bar does not bend).

This pattern of results corresponds to the rule affirmation of p and can be stated as long bars bend more than short bars.

Question 6 in the bending rods problem states the hypothesis that brass bars bend more than steel bars and asks for the necessary conditions for this hypothesis to be true and for the conditions which must not exist for the hypothesis to be true. The use of a particular logical rule to answer the question can be assessed by the two types of conditions stated by the subject. This particular hypothesis, for example, could be interpreted as either implication, exclusive disjunction or affirmation of q. To illustrate different interpretations, the implication interpretation and the affirmation interpretation will be contrasted. If subjects conform to an implication rule, the following statements would be given:

pq 1 (the long brass bar should bend)

$p\bar{q}$ 0 (the long steel bar should not bend)

$\bar{p}q$ 1 (the short brass bar should bend)

$\bar{p}\bar{q}$ 1 (the short steel bar should bend).

If, on the other hand, subjects conform to an affirmation of q rule, the following statements would be given:

pq 1 (the long brass bar should bend)

$p\bar{q}$ 0 (the long steel bar should not bend)

$\bar{p}q$ 1 (the short brass bar should bend)

$\bar{p}\bar{q}$ 0 (the short steel bar should not bend).

Question 7 asks for the opposite or reverse hypothesis from the one stated in the previous question. If subjects interpret "reversal" as Piaget does (in terms of the INRC group) and used an implication rule in Question 6, according to Brainerd (1978, see Table 6-2), they should answer in terms of reverse implication:

pq 1 (the long brass bar should bend)

$p\bar{q}$ 1 (the long steel bar should bend)

$\bar{p}q$ 0 (the short brass bar should not bend)

$\bar{p}\bar{q}$ 1 (the short steel bar should bend).

If affirmation of q was the rule used in Question 6, subjects should answer in terms of denial of q:

pq 0 (the long brass bar should not bend)

$p\bar{q}$ 1 (the long steel bar should bend)

$\bar{p}q$ 0 (the short brass bar should not bend)

$\bar{p}\bar{q}$ 1 (the short steel bar should bend).

In Question 8, the subject is told that both the 20-inch bars bend down one inch and both of the 10-inch bars do not bend. The subject is asked to state a conclusion based on these results. The expected conclusion is that long bars

bend more than short bars. Subjects are also asked to state what results or results would disprove their conclusion if these results existed. For example, if affirmation of p is the interpretation of this conclusion, the existence of the $\bar{p}q$ case (the short brass bar bends) and the $\bar{p}\bar{q}$ case (the short steel bar bends) would be disproving.

Question 9 also states specific results, and the subject is asked to form a conclusion. It is stated that the only bar that bends is the 10-inch steel bar. These results represent an example of a logically possible case (conjunctive negation) that is not congruent with everyday experience or knowledge. If a subject answers on the basis of this logical rule, the conclusion would be stated as neither length nor material have an effect on the bending of the bars.

The final question in the bending rods problem states the hypothesis long bars bend if and only if they are brass, and the subject is asked to state which results must exist and must not exist for this hypothesis to be true. Two possibilities for the logical rule interpretation of this problem are biconditional and conjunction. If subjects conform to a biconditional rule, the following statements would be given:

pq 1 (the long brass bar should bend)

$p\bar{q}$ 0 (the long steel bar should not bend)

$\bar{p}q$ 0 (the short brass bar should not bend)

$\bar{p}\bar{q}$ 1 (the short steel bar should bend).

However, if subjects conform to a conjunction rule, the following statements would be given

pq 1 (the long brass bar should bend)

$p\bar{q}$ 0 (the long steel bar should not bend)

$\bar{p}q$ 0 (the short brass bar should not bend)

$\bar{p}\bar{q}$ 0 (the short steel bar should not bend)

In the pendulum problem, the last four questions are designed to assess the use or nonuse of propositional rules as well as the use of control of variables. In these questions, the hypotheses and results are stated such that p and q can refer to the dependent and independent variables, respectively (see Dawes, 1972). Since this problem involved more variables that should be held constant than the bending rods problem, only nine questions were asked so that the subject could complete the problem in approximately an hour. The questions are very similar in logical structure to the bending rods questions, and thus will not be discussed in detail (see Appendix E).

In addition, the information and instruction sheet utilized in Experiment 1 was modified for use in this experiment. The same information concerning laboratories and classes was asked of the subject, but subjects were only instructed that they would be given two problems to complete. This information sheet is given in Appendix E. The subjects were given specific instructions on the first page of each problem. They were told to answer each question as exactly

and completely as possible and were told to answer each question in order.

Procedure

Subjects were run in same-sex groups of from eight to 10 subjects each in two one-hour sessions separated by one day. The subjects completed the information sheet on the first day, and then were given one of the problems to solve. The order of presentation of the two problems was counterbalanced within sex. Subjects were told that the problems were self-paced and that they would be allowed as much time as necessary to complete the tasks. They were also reminded to answer each question as exactly and completely as possible and to ask the experimenter if they had any questions concerning the task.

Results

The average age of the subjects was 19.65 years (SD= 3.21). Since the classes taken in Experiment 1 were so highly intercorrelated, the total number of classes taken for each subject was calculated and used in all further analyses. The mean total number of classes taken for these subjects was 10.19 (SD=5.15). The greatest number of classes taken was 28, and the fewest was two. A total SAT score was also calculated and used in all further analyses because of the high correlation between SAT (V) and SAT (Q) in Experiment 1.

The mean total SAT score was 1034.39 points (SD=147.16). The highest actual score was 1370 points and the lowest actual score was 660 points.

The Bending Rods Problem

The first question was scored by counting the number of factors listed that could affect the bending of the bars. It was anticipated that three factors would be listed by most subjects because these factors were the most obvious ones to list from reading the description of the problem. These three factors were the length of the bar, the material of the bar, and the amount of weight attached to the end of the bar. The mean number of these expected factors listed was 2.37 (SD=.88). The number of additional factors was also counted. These factors included such variables as which position the bar is in, the atmospheric conditions of the test (e.g., normal versus in a vacuum), and where the weight was placed on the bar. The mean number of additional factors reported was 1.76 (SD=1.71).

Questions 2 through 5 were scored as either correct (1) or incorrect (0). Thirty-nine of the 46 subjects (85%) answered all four of these questions correctly. Five subjects missed one question, and two subjects missed two questions. The mean percent correct for Questions 2 through 5 were 93%, 98%, 96% and 93%, respectively. Forty-five of the 46 subjects answered Question 6 in a manner congruent with the

logical rule affirmation of q. These subjects stated that both the 10-inch and 20-inch brass bars must bend, and the 10-inch and 20-inch steel bars must not bend, or must bend less than the brass bars. The remaining subject answered in terms of length rather than material.

The same 45 subjects answered Question 7 in terms of denial of q, but the rule was actually stated in terms of affirmation of $\bar{q}$, which is an equivalent expression: steel bars bend more than brass bars. This reverse hypothesis is the reverse hypothesis of affirmation of q in terms of the INRC group (Brainerd, 1978, Table 6-2). Subjects answering this question correctly stated that both lengths of steel must bend and both lengths of brass must not bend or must bend less than the steel bars. One of these subjects expressed reservations about the hypothesis based on her past experience, an indication of a concrete-operational level of behavior. The same subject who missed Question 6 missed this question by answering "same as Number 6."

Question 8 presented results and asked subjects for a conclusion from these results. The answers fell into several categories. These categories are presented in Table 11 along with the number of subjects answering in each category. Most of the subjects (25 of 46 subjects) gave the expected answer that length was the crucial factor for bending of the bars. Fifteen subjects made an inference about the results and stated that brass and steel had equal strength; three of

Table 11.

The Number of Subjects for Each Type of Answer Given to Question 8 on the Bending Rods Problem (Experiment 2)

	Answer	N
1.	Length is the crucial variable	25
2.	The strengths of brass and steel are equal	12
3.	Both 1 and 2	3
4.	The experiment was confounded	2
5.	Either 1 or 4	1
6.	The results are not possible	1
7.	Nothing can be concluded	2
	Total	46

these subjects also stated that length is a factor in bending. This inference about the results can be interpreted as a higher-level answer than stating that length is the crucial variable (Peel, 1971). The remaining six subjects gave various answers: two subjects stated that the experiment must have been confounded in some way, one subject did not accept the results, two subjects stated that no conclusion could be made without further information, and one subject stated that either length was the crucial factor or the experiment was confounded.

Question 9 also presented results to be evaluated, but these results are based on a logical rule which does not have an experiential basis. Answers to this question again formed several categories, which are given in Table 12 along with the number of subjects who gave each answer. No subject gave an answer that conformed to the logical rule of <u>conjunctive negation</u> that the results were based on. Half of the subjects (23 of 46) gave a restricted answer; namely that the short steel bar bends. This answer is less than formal operational because it is restricted to the specific concrete evidence at hand (e.g., Flavell, 1963; Peel, 1971). Thirteen subjects stated that the experiment must have been confounded, and five subjects stated that nothing could be concluded from the results without further information. One subject stated that the results were not possible, and four subjects missed the question by leaving it blank.

Table 12.

The Number of Subjects for Each Type of Answer Given to Question 9 on the Bending Rods Problem (Experiment 2)

Answer	N
1. The short steel bar bends	23
2. The experiment was confounded	13
3. Nothing can be concluded	5
4. Missed the question	4
5. The results are not possible	1
Total	46

All subjects answered question 10 in accordance with the logical rule of <u>conjunction</u> by stating that only the long brass bar should bend. Four of the subjects stated that all possible metals must be tested.

The Pendulum Problem

The first question was scored as in the bending rods problem. It was anticipated that four factors would be listed by most subjects since these factors were stated in the description of the problem. These four factors were the length of the string, the amount of weight attached to the end of the string, the height of the dropping point, and the force behind the initial swing. The mean number of factors listed for these expected factors was 3.22 (SD=.81). The number of additional factors reported was 1.87 (SD=1.89).

Questions 2 through 5 were scored as correct (1) or incorrect (0). Thirty-eight of the 46 subjects (82.6%) answered all four of these questions correctly. Four subjects missed one question, one subject missed two questions, one subject missed three questions, and two subjects missed all four questions. The mean percents correct for Questions 2 through 5 were 89%, 91%, 91% and 91%, respectively.

Question 6 asked subjects how to test a specific hypothesis, and which results must exist and not exist if the hypothesis was true. Thirty-eight of the 46 subjects answered the question correctly by stating that they would hold the

pendulum at different heights and time the number of swings that occurred for each height. In addition, they stated that pendulums held higher must swing faster than pendulums held lower if the hypothesis was true. These subjects differed in the number of factors that they stated which needed to be controlled. The factors which were expected to be listed as controlling variables were the length of the string, the amount of weight added to the string, and the initial force of propulsion. Of the 38 subjects who answered correctly, eight subjects mentioned no controlling factors, two subjects mentioned only one factor, 13 subjects mentioned length of string and weight as factors to be controlled, and 15 subjects mentioned all three factors. The eight subjects who missed this question did so because either their answers were not specific enough to be scored as correct or they stated that height would not matter and thus did not answer the question.

Question 7 presents results and probes specifically for mention of controlling variables by asking if any conditions must be met before the results are interpretable. Fourteen subjects stated that no conditions had to be met, 14 subjects mentioned one controlling variable, and 17 subjects mentioned two controlling variables. One subject did not answer this question. Of the 45 subjects who did answer the question, 44 correctly interpreted the results as demonstrating length as the crucial variable. One subject gave no answer to this part of the question.

A hypothesis was presented in Question 8, and the subject's task was to state how he or she would test the hypothesis and what results must exist and not exist for the hypothesis to be true. Thirty-eight of 45 subjects (one subject did not answer the question) or 84.4% of the subjects answered correctly by stating that they would test each weight and time the resulting number of oscillations per minute. Again, subjects differed in the number of controlling variables listed--eight subjects listed no controlling variables, six subjects listed one, eight subjects listed two, and 16 subjects listed three controlling variables. Six of the subjects who answered this question incorrectly also answered Question 6 incorrectly. Answers were judged as incorrect if the subject stated that weight did not affect oscillation, or if their answers were too general to be scored as correct.

Forty-one of the forty-six subjects (89%) answered the last question for the pendulum problem correctly by stating that short strings result in fewer swings per minute, with weight not effecting the rate of oscillation. Three of these subjects mentioned controlling the height of release, and two subjects mentioned controlling for force of propulsion as well. Five subjects answered incorrectly and one subject did not answer this question.

Summary

To summarize the data so far, the majority of subjects do conform to logical rules when questions are specific enough to elicit this behavior Thus, it is possible for subjects to be logically consistent in answering questions and to evaluate hypotheses and results in terms of propositional logic, at least for the questions formulated for this experiment. However, based on the results of Experiment 1, subjects probably do not use this type of logic for open-ended questions utilized in the traditional Piagetian tasks. The use of isolation and control of variables along with the use of combinations appears to be the most important behaviors for successful performance on traditional Piagetian tasks.

The majority of subjects in this experiment also were able to isolate variables and easily recognize confounded comparisons of variables. Subjects mentioned the control of variables with a much wider range of responses. This may be the result of the written nature of the task. As in Experiment 1, it is possible that some subjects simply were not motivated to write down all of the variables that it is truly necessary to control in a particular experiment.

Analysis investigating the relationship between the factors

The relationship between the measured factors and performance on the questions was investigated using Pearson

moment-product correlations. The resulting correlation matrix is given in Table 11. These correlations will be discussed in detail because the number of subjects is not sufficient for a partial correlation analysis. Only significant correlations ($p<.05$) will be discussed; therefore, 'correlated' will henceforth refer to significantly correlated. Both the age and the number of classes taken are correlated with year in school (r=.37 and .50, respectively), which are not surprising results. Sex is correlated with SAT score, (r=-.40) with males scoring higher than females. Sex is also correlated with the number of factors given in the pendulum problem on Question 1 (r=-.34), with males again with a superior score.

The number of classes taken is correlated with SAT score (r=.50), and also with several questions in the pendulum problem--the two isolation of variables questions (r=.31 and .32 for Questions 2 and 3) and Questions 7 and 8 (r=.41 and .33, respectively). Most subjects answered Questions 6 through 9 correctly, and performance is distinguished primarily by the number of controlling variables mentioned by the subject. Thus, subjects who score higher on the SAT exam and have taken more classes tend to mention more controlling variables. In support of this hypothesis, SAT is also correlated with the number of factors listed in both problems (r=.33 for the bending rods problem and =.57 for the pendulum problem) and with the last four questions in

Table 13.

The Correlation Matrix for the Factors in Experiment 2

	AGE	SEX	YR	CL	SAT	FAC	XFAC	IS1
AGE	1.00	.14	.37*	-.01	.08	.11	-.15	-.001
SEX		1.00	-.01	-.14	-.40*	-.04	-.16	-.08
YR			1.00	.50*	.20	.07	-.04	.01
CL				1.00	.50*	-.03	.21	.13
SAT					1.00	.39*	.20	.24
FAC						1.00	-.18	-.09
XFAC							1.00	-.09
IS1								1.00
IS2								
CON1								
CON2								
SIX								
SEV								
EIG								
NIN								
TEN								
PFAC								
PXFAC								
PIS1								
PIS2								
PCON1								
PCON2								
PSIX								
PSEV								
PEIG								
PNIN								

*$p < .05$

Legend: AGE = age of subject, SEX = sex of subject, YR = year in school, CL = total number of measured classes taken, SAT = combined SAT (V) and SAT (Q) scores, FAC = number of expected factors listed in the bending rods problem, XFAC = number of additional factors listed in the bending rods problem, IS1 = the first question in the bending rods problem that tests for knowledge of isolation of variables, IS2 = the second question in the bending rods problem that tests for knowledge of isolation of variables, CON1 = the first question in the bending rods problem that tests for recognition of the confounding of variables, CON2 = the second question in the bending rods problem that tests for recognition of

Table 13--continued

	IS2	CON1	CON2	SIX	SEV	EIG	NIN	TEN
AGE	.03	.08	.08	.08	.08	06	-.06	-.01
SEX	-.14	.01	-.08	-.14	-.14	-.13	-.16	.14
YR	.11	-.18	.01	.11	.11	.10	.21	.01
CL	-.17	-.20	.15	.06	.06	.12	.24	-.09
SAT	.05	.05	.14	.00	.00	.03	.08	-.12
FAC	.24	-.03	-.09	-.11	-.11	-.32*	-.09	.05
XFAC	-.02	.03	.12	.07	.07	.07	.30*	-.14
IS1	-.04	-.06	.29	-.04	-.04	.12	-.11	.08
IS2	1.00	-.03	-.04	-.02	-.02	-.02	.11	.05
CON1		1.00	.38*	-.03	-.03	.07	-.06	07
CON2			1.00	-.04	-.04	.20	.06	.08
SIX				1.00	1.00*	.11	.19	.05
SEV					1.00	.11	.19	.05
EIG						1.00	.23	.11
NIN							1.00	.28
TEN								1.00
PFAC								
PXFAC								
PIS1								
PIS2								
PCON1								
PCON2								
PSIX								
PSEV								
PEIG								
PNIN								

the confounding of variables, SIX = question six in the bending rods problem, SEV = question seven in the bending rods problem, EIG = question eight in the bending rods problem, NIN = question nine in the bending rods problem, TEN = question ten in the bending rods problem, PFAC = number of expected factors listed in the pendulum problem, PXFAC = number of additional factors listed in the pendulum problem, PIS1 = the first question in the pendulum problem that tests for isolation of variables, PIS2 = the second question in the pendulum problem that tests for isolation of variables, PCON1 = the first question in the pendulum problem that tests for recognition of the confounding

Table 13--continued.

	PFAC	PXFAC	PIS1	PIS2	PCON1	PCON2	PSIX	PSEV
AGE	.14	-.16	-.05	-.02	.04	.01	.04	-.14
SEX	-.34*	-.19	-.04	-.20	.01	.14	-.21	-.14
YR	.08	-.25	.24	.19	-.09	.09	.09	.15
CL	.20	-.02	.31*	.32*	.03	.07	.22	.41*
SAT	.57*	.09	.24	.29	.12	.27	.59*	.43*
FAC	.23	-.04	.06	.21	.31*	.22	.24	.27
SFAC	.28	.57*	.13	.03	.14	.14	.27	.23
IS1	.07	.02	.17	.18	-.08	-.08	.07	.13
IS2	.04	.01	-.13	.03	-.05	-.05	-.16	-.17
CON1	.06	.13	-.06	-.12	.31*	.31*	.12	.02
CON2	.29	.07	.17	.18	.23	.23	.13	.24
SIX	.23	.09	.04	.03	-.05	-.05	.23	.19
SEV	.23	.09	.04	.03	-.05	-.05	.23	.19
EIG	.13	.25	.24	.28	-.04	.03	.02	-.08
NIN	.08	.03	.19	.23	.07	-.08	.19	.02
TEN	-.18	-.07	.003	.05	.10	.10	.12	.16
PFAC	1.00	.02	.27	.31*	.28	.37*	.57*	.28
PXFAC		1.00	-.07	.07	.19	.15	.10	-.04
PIS1			1.00	.76*	.26	.35*	.52*	.43*
PIS2				1.00	.29	.40*	.48*	.42*
PCON1					1.00	.45*	.28	.21
PCON2						1.00	.48*	.31*
PSIX							1.00	.65*
PSEV								1.00
PEIG								
PNIN								

of variables, PCON2 = the second question in the pendulum problem that tests for recognition of the confounding of variables, PSIX = question six in the pendulum problem, PSEV = question seven in the pendulum problem, PEIG = question eight in the pendulum problem, PNIN = question nine in the pendulum problem.

Table 13—continued.

	PEIG	PNIN
AGE	.14	.06
SEX	-.19	-.13
YR	.18	.06
CL	.39*	.16
SAT	.66*	.49*
FAC	.32*	.27
XFAC	.002	.20
IS1	.29	.16
IS2	.13	.01
CON1	.05	.01
CON2	.24	.01
SIX	.23	.01
SEV	.23	.01
EIG	-.01	.05
NIN	.14	.32
TEN	.23	.13
PFAC	.39*	.20
PXFAC	-.07	-.11
PIS1	.33*	.14
PIS2	.45*	.19
PCON1	.12	.16
PCON2	.43*	.16
PSIX	.63*	.43*
PSEV	.52*	.25
PEIG	1.00	.37*
PNIN		1.00

the pendulum problem (r=.59, .43, .66 and .49 for Questions 6 through 9, respectively). In addition, the number of factors listed in the pendulum problem is correlated with questions in the pendulum problem involving isolation of variables (r=.31), control of variables (r=.37), and Questions 6 (r=.57) and 8 (r=.39). This pattern of correlations suggests that subjects who list relevant factors in the first question are more likely to control those factors in later questions. The questions within the pendulum problem are highly intercorrelated as well, which reflects the primary basis on which subjects' performance was scored--control of variables.

Performance on the bending rods questions does not appear to be as intercorrelated. The number of factors listed is correlated with performance on Question 8 (r=-.32). This negative correlation reflects the tendency for subjects who listed the most factors to state that length was the relevant factor in bending based on the given results. Subjects who listed fewer factors tended to state that the experiment was confounded, refused to accept the results, or stated that nothing could be concluded without further information.

The number of additional factors listed is correlated with performance on Question 9 (r=.30). This correlation reflects the tendency for subjects who listed fewer additional factors to state the restricted answer that the short steel bar bends. Subjects who listed more additional factors

stated that the experiment was confounded, refused to accept the results, or stated that more information was needed before a conclusion could be drawn.

Performance on the two questions that were designed to probe for recognition of confounded variables was correlated as well ($r = .38$). Also, performance on Questions 6 and 7 was perfectly correlated ($r = 1.00$). This reflects the fact that all subjects who conformed to the affirmation of q rule for Question 6 stated that the reverse hypothesis would be denial of q.

The analysis for the relationships between the factors in this experiment provides further support for the hypothesis stated earlier in this section. Subjects appear to be able to answer questions consistently based on both formal operational abilities and propositional logic. However, in the bending rods problem, the two types of questions were not related to a great extent, and the relationship between the two types of questions in the pendulum problem is due to the use of controlling variables in all questions. Subjects answered consistently in terms of propositional logic rules, but the interpretation of results for the more complex questions fell into several different categories. In addition, the propositional rules that subjects conformed to were usually the simplest interpretation of the results.

CHAPTER 7
GENERAL DISCUSSION AND CONCLUSIONS

The results of Experiment 1 revealed no relationship between performance on pencil-and-paper versions of traditional Piagetian tasks and on the THOG problem and selection task. This result was interpreted as an indication that different abilities may underlie performance on the two types of tasks. Specifically, it was inferred that the most important ability for successful performance on the Piagetian tasks was the isolation and control of variables. It was inferred that successful performance on the two logic tasks was due to propositional logic ability.

However, alternative interpretations for these results are possible. For successful performance on the selection task, subjects must consider each card individually, form a hypothesis about what could be on the reverse side, and evaluate the consequences of the alternatives. For successful performance on the THOG task, subjects must postulate hypotheses about what must be written down based on the rule and the instance of a THOG, separate the attributes of each design from the design itself, perform a combinatorial analysis on the resulting hypothesized relevant attributes, and finally, form the correct conclusion based upon the results of the preceeding three steps in the problem. The Piagetian

tasks require the subject to postulate possible causes of the variable of interest and then exhaustively test these possible causes in a controlled experiment. Performance on the Piagetian tasks was much higher than that on the propositional tasks in Experiment 1. Thus, although the propositional logic tasks are presented within a problem-solving framework which makes the overall structure of all of the tasks very similar, this framework also creates additional steps for the solution of the propositional logic problems. The differences in task performance, therefore, could be due to task differences rather than differences in underlying ability; the propositional logic tasks appear to be more difficult.

The data of Experiment 1 are not sufficient to separate these two possible explanations. However, Lawson et al. (1978) tested different age groups utilizing Piagetian tasks and a version of Wason's selection task. When a principle-components analysis of the data was performed by grade level, the tasks did not load according to level of difficulty at any grade level. Lawson et al. (1978) interpret this result as supportive of the hypothesis that different abilities underlie performance on the Piagetian and propositional logic tasks rather than task differences.

However, a further difficulty exists in this interpretation of the data of Experiment 1. A Piagetian interpretation of the results would explain the performance differences in

terms of the abstractness of the propositional logic tasks — they are outside of the realm of anyone's experience, therefore it is not surprising that performance level is so low. As was discussed in Chapter 3, however, merely presenting these propositional logic tasks in a realistic framework does not improve performance. Only if a specific memory-cuing situation is utilized will a majority of subjects give the correct answer. It is possible that some transfer exists between a specific situation and a similar situation-based rule, possibly by using a reasoning-by-analogy strategy (e.g., Cox & Griggs, Note 6; D'Andrade, cited in Rumelhart, 1980; Golding, 1981). The distinction between the use of logic and the use of memory, then, becomes blurred. It is not clear how to resolve this paradox for either the logic tasks or the Piagetian tasks. To test the relationship between a subject's knowledge base and his or her use of propositional logic, the knowledge framework will have to be evoked by the problem without eliciting the specific solution to the problem. One approach to this strategy has been to formulate the task so that the correct answer is contradictory to experience (e.g., Newstead et al., Note 2, Experiment 4), and there may be other approaches.

Because of these considerations, any conclusions about the nonsignificant correlation between the Piagetian tasks and the propositional logic tasks given in Experiment 1 must be qualified. However, it is clear that successful

performance on the propositional logic tasks is not a prerequisite for successful performance on the Piagetian tasks, and Lawson et al. (1978) present evidence that separate abilities underlie the two types of tasks.

Subjects in Experiment 2 performed at a very high level. Almost every subject listed at least some factors which could influence both presented situations. In addition, most subjects recognized confounded variables and could isolate and test variables when asked to do so. The greatest variability in performance occurred with the number of controlling factors that subjects mentioned. Again, this supports the hypothesis that a major component of formal operational behavior is the ability to isolate and control variables.

Subjects also conformed to propositional logic rules in interpreting many of the presented results and in reporting the necessary results to support a particular hypothesis. The logic rule that was used by the subject, however, was always the simplest one possible. In addition, when the presented results were based on a rule that was not congruent with everyday experience, a variety of qualitatively different response categories occurred. Several categories represented concrete operational responses since the answers were more restricted (Peel, 1971). However, this situation occurred on only a few questions, and for the most part subjects performed at a formal operational level. One limitation of the questions utilized in Experiment 2 is that a

wider range of questions representing more of the propositional logic rules probably would have differentiated between performance levels to a greater extent

A comparison between the performance levels of the subjects in Experiments 1 and 2 reveals that subjects performed on the average at a higher level in Experiment 2. In other words, subjects perform better when the questions are more specific, when they are given a chance to recognize the confounding of factors, and when they are asked specific questions about how to test for the effects of a specific factor. Also, subjects do conform to propositional logic if a question poses a problem in such a way that this conformity can be easily measured. Thus, it is possible to conclude that most people in this age range do perform at the formal operational level, and in accordance with propositional logic, but that this ability is not elicited by the typical formal operational task.

However, this conclusion may be premature because of the question of how formal operations should be defined. To define formal operations as the ability to recognize certain erroneous aspects of data or the ability to answer specific questions in a manner consistent with propositional logic may not be as valid as a definition employing the idea of the spontaneous use of this knowledge. This argument is similar to that expressed by Keating (1980) questioning the technique of assessing formal operations by mere permutational ability.

[illegible] consistent with the original formulation of formal operations by Piaget to assume that a truly formal operational person will interpret reality in terms of an overriding framework congruent with formal operations, and thus will be able to express this ability in many appropriate situations.

This framework hypothesis can be conceptualized without reference to a core structure consisting of the 16 binary propositional operations. In several studies, Lawson and his colleagues and others have demonstrated that the abilities typically associated with formal operations, such as correlations, probability, proportional reasoning, combinatorial analyses, and the isolation and control of variables develop during adolescence (e.g., Kuhn & Brannock, 1977; Lawson et al., 1978, Lawson & Nordland, 1976; Lawson & Renner, 1974). Lawson et al. (1978) conclude that the advances in hypothetico-deductive reasoning are the result of these developing "schemes," and that propositional logic plays essentially no role in this process.

This conclusion postulating the relevance of these schemes can account for most of the results in past research on formal operations. In fact, researchers have not usually attempted to link the results of their experiments to the 16 binary operations, but rather have usually related their findings to one or more of the formal operational abilities. The present research, and that of Lawson and his colleagues, suggests that this connection is probably not necessary.

APPENDIX A
THE FOUR CARD PROBLEM AND JUSTIFICATION
SHEET USED IN EXPERIMENT 1

A	B	2	3
Card 1	Card 2	Card 3	Card 4

The cards above have information on both sides. On one side of a card is a letter, and on the other side is a number.

Here is a rule:

IF A CARD HAS A VOWEL ON ONE SIDE, THEN IT HAS AN EVEN NUMBER ON THE OTHER SIDE.

Select those cards that you definitely need to turn over to determine whether or not the four cards are violating the rule.

ANSWER ______________________________

Now that you have indicated your answers, would you please give your reasons for the responses you chose? Please write, in the space below, why you chose the card or cards that you did.

Please note that you <u>MUST NOT CHANGE</u> the answers you have already given. If while writing your reasons you decide that your answer was wrong, you may say so in the space below but leave the original response sheet unchanged.

<u>Card Number</u> <u>Reasons you chose the card</u>

1

2

3

4

APPENDIX B
THE THOG PROBLEM AND JUSTIFICATION SHEET
USED IN EXPERIMENT 1

At the bottom of the page are four designs BLACK DIAMOND, WHITE DIAMOND, BLACK CIRCLE, and WHITE CIRCLE. You are to assume that I have written down one of the colors (black or white) and one of the shapes (diamond or circle). Now read the following rule carefully.

IF A DESIGN INCLUDES EITHER THE COLOR THAT I HAVE WRITTEN DOWN OR THE SHAPE THAT I HAVE WRITTEN DOWN, THEN IT IS CALLED A THOG. IF IT INCLUDES <u>BOTH</u> THE COLOR AND THE SHAPE THAT I HAVE WRITTEN DOWN OR <u>NEITHER</u> THE COLOR OR THE SHAPE THAT I HAVE WRITTEN DOWN, THEN IT IS <u>NOT</u> CALLED A THOG.

I now tell you that the BLACK DIAMOND is a THOG. Given this information, each of the other designs can now be classified into one of the following categories:

A) Definitely is a THOG
B) Definitely is <u>not</u> a THOG
C) Insufficient information given to decide if a THOG.

Using the rule and the information that the BLACK DIAMOND is a THOG, please classify each of the other three patterns in terms of A, B, and C.

Classification A ______ ______ ______

Now that you have indicated your answers, would you please give your reasons for the responses you chose? Please write, in the space below, why you categorized each design in the way you did.

Please note that you MUST NOT CHANGE the answers you have already given. If while writing your reasons, you decide that your answer was wrong, you may say so in the space below but leave the original response sheet unchanged

Design Reasons for choice of classification

white diamond

black circle

white circle

APPENDIX C
THE PROBLEM BOOKLET USED IN
EXPERIMENT 1

Name________________Sex_____Age_____Classification_____

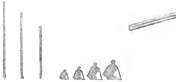
Please list the courses you have had in

Math High School	Lab?	College	Lab?
Chemistry High School	Lab?	College	Lab?
Physics High School	Lab?	College	Lab?
Logic High School	Lab?	College	Lab?

You will be asked to solve three problems in this part of the experiment. I am not asking you to demonstrate your knowledge about the answers to the problems. Rather, I am interested in your method of solving the problems. Therefore, whether or not you already know the answer from your past experience, it is very important that you explain the method for finding the answer carefully and completely. When you answer the question, you can assume that you have unlimited materials of any type. It is also important that you write clearly so that I can read your answers!

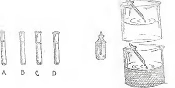
This drawing represents a wooden block with several openings. In the openings, metal bars can be inserted, as drawn here. Also, weights can be placed at the end of the bars. The weights will bend these metal bars. Suppose you are given four bars. Two bars are brass, and two bars are steel. Of the two brass bars, one is 10-inches long and one is 20-inches long. Of the two steel bars, one is also 10-inches long and the other 20-inches long.

The idea behind this problem is to design a way to find out what factors influence how much the bars bend under the weight. Please describe as exactly and as completely as possible the way you would go about finding out what factors influence the bending of the bar. Remember, whether or not you already know the answer from your past experience, it is very important that you explain the method for finding the answer carefully and completely.



This drawing represents a set of strings of various lengths and a set of varying weights. A string can be attached to the overhanging bar and a weight can be attached to the string to form a pendulum. The weight at the end of the string can be lifted to various heights and dropped to start the swinging motion. The weight can also be propelled with varying forces when dropped.

The idea behind this problem is to design a way to find out what factors influence how fast the pendulum swings back and forth. Please describe as exactly and as completely as possible the way you would go about finding out what factors influence the speed of oscillation of the pendulum. Remember, whether or not you already know the answer from your past experience, it is very important that you explain the method for finding the answer carefully and completely.



This drawing represents a set of four test tubes which contain colorless, odorless chemicals (in other words, they are perceptually identical). The dropper bottle also contains a chemical. Suppose you are shown two beakers which contain perceptually identical liquids which are made up of some or all of the chemicals in the test tubes (A, B, C, D). When the substance in the dropper bottle is added to one of the beakers, nothing happens. When it is added to the other beaker, a yellow color appears.

The idea behind this problem is to design a way to reproduce the liquid that turns yellow when the substance in the dropper bottle is added. Please describe as exactly and completely as possible the way you would go about finding out what makes a yellow color appear. Remember, whether or not you already know the answer from your past experience, it is very important that you explain the method for finding the answer carefully and completely.

APPENDIX D

THE PERMUTATION PROBLEM PRESENTED IN EXPERIMENT 1

This drawing represents a license plate with four digits. The idea behind this problem is to write down all of the possible four digit license plate numbers which use each of these digits exactly once.

Can you figure out a general rule which would tell you how many licenses you would have for any number of digits (n)?

APPENDIX E
THE INFORMATION SHEET AND TWO PROBLEMS USED IN EXPERIMENT 2

Name__________________ Sex______ Age______ Classification______

Please list the courses you have had in

Math
High School Lab? College Lab?

Chemistry
High School Lab? College Lab?

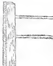
Physics
High School Lab? College Lab?

Logic
High School Lab? College Lab?

You will be asked questions about two problems in this experiment.[a] Please write clearly so that I can read your answer!

---

[a]In the experiment, only two questions were on each page. For reasons of brevity, more than two questions per page are presented in this Appendix.



This drawing represents a wooden block with several openings. In the openings, metal bars can be inserted, as drawn here. Also, weights can be placed at the end of the bars. The weights will bend these metal bars. Suppose you are given four bars. Two bars are brass, and two bars are steel. Of the two brass bars, one is 10-inches long and one is 20-inches long. Of the two steel bars, one is also 10-inches long and the other 20-inches long.

Please answer the following questions as exactly and as completely as possible. Complete each question before going on and do not go back to previous questions. If you decide to modify an answer while on another question, talk to the experimenter.

1. List all of the possible factors which could effect the bending of the bars.

2. Which two bars can you compare to find out if long bars bend more than short bars? Why did you choose these two bars to compare? Are there any other comparisons that would be helpful for this question?

3. Which two bars would you compare to find out if brass bars bend more than steel bars? Why did you choose these two bars to compare? Are there any other comparisons that would be helpful for this question?

4. Could you compare the 10-inch steel bar with the 20-inch brass bar to see if long bars bend more than short bars? Why or why not?

5. Could you compare the 10-inch brass bar with the 20-inch steel bar to see if brass bars bend more than steel bars? Why or why not?

6. Suppose you wanted to test the hypothesis that brass bars bend more than steel bars. If this statement is true, what result or results in the experiment must exist? What result or results must not exist? (In other words, which bar or bars should bend and which should not bend?)

7. Look at the hypothesis stated in Question 6. What would be the opposite (or reverse) hypothesis? What result or results must exist for this hypothesis to be true? What result or results must not exist? (What bar or bars should bend and which should not bend?)

8. Suppose that you observed the following results:

20-inch brass bar:	bends down 1 inch
20-inch steel bar:	bends down 1 inch
10-inch brass bar:	does not bend
10-inch steel bar:	does not bend

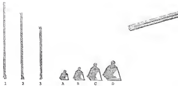
What could you conclude from these results (if anything)? What result or results in the experiment would disprove your conclusion if these results existed?

9. Suppose that you observed the following results:

20-inch brass bar:	does not bend
20-inch steel bar:	does not bend
10-inch brass bar:	does not bend
10-inch steel bar:	bends down 1 inch

What could you conclude from these results (if anything)? What result or results in the experiment would disprove your conclusion if they existed?

16. Suppose that you wanted to test the hypothesis that long bars bend if and only if they are brass. If this statement is true, what result or results in the experiment must exist? (Which bar or bars must bend?) Which result or results would disprove the hypothesis if they did exist? (Which bar or bars must not bend?)



This drawing represents a set of strings of various lengths and a set of varying weights. A string can be attached to the overhanging bar and a weight can be attached to the string to form a pendulum. The weight at the end of the string can be lifted to various heights and dropped to start the swinging motion. The weight can also be propelled with varying forces when dropped.

Please answer the following questions as exactly and as completely as possible. Complete each question before going on and do not go back to previous questions. If you decide to modify an answer while on another question, talk to the experimenter.

1 List all of the possible factors which could affect the speed of oscillation of the pendulum.

2. Which strings along with which weights would you compare to find out if there are more swings per minute for longer strings or for shorter strings? Why did you choose these particular combinations? Are there any other comparisons that would be helpful for this question?

3. Which strings along with which weights would you compare to find out if there are more swings per minute for heavier weights or for lighter weights? Why did you choose these particular combinations? Are there any other comparisons that would be helpful for this question?

4. Could you compare string 1-weight A with string 3-weight B to see if there are more swings per minute for longer strings or for shorter strings? Why or why not?

5. Could you compare string 1-weight D with string 2-weight B to see if there are more swings per minute for heavier weights or for lighter weights? Why or why not?

6. Suppose that you wanted to test the hypothesis that the higher you held the weight before releasing it, the more swings per minute would result. Now, specifically, would you test this hypothesis? If this statement is true, what specific results must exist? What results must not exist?

7 Suppose that you observed the following results:

string 1-weight C least number of swings per minute
string 2-weight C. medium number of swings per minute
string 3-weight C greatest number of swings per minute

Are there any specific conditions that must be met in this experiment before the above results are interpretable? What could you conclude from these results (if anything)? What result or results in the experiment would disprove your conclusion if they existed?

8. Suppose that you wanted to test the hypothesis that heavier weights result in more swings per minute. How, specifically, would you test this hypothesis? If this statement is true, what specific results must exist? What results must not exist?

9. Suppose that you observed the following results:

string 1-weight A: few swings per minute
string 3-weight A: many swings per minute
string 1-weight B: few swings per minute
string 3-weight B: many swings per minute

What could you conclude from these results (if anything)? What result or results would disprove your conclusion if these results existed?

REFERENCE NOTES

1. Neimark, E. D. Confounding with cognitive style factors: An artifact explanation for the apparent non-universal incidence of formal operations. Unpublished manuscript, Douglass College, Rutgers, undated.

2. Newstead, S. E., Griggs, R. A., & Warner, S. A. The effects of realism on Wason's THOG problem. Manuscript submitted for publication, June 1981.

3. Griggs, R. A., & Newstead, S. E. The effects of problem-phrasing and memory-cuing on performance on Wason's THOG problem. Unpublished manuscript, University of Florida, 1981.

4. Pauet, D., & Golembowski, G. The validation of a group administered Piagetian measure. Unpublished manuscript, Ohio University, undated.

5. Tomlinson-Keasey, C., & Campbell, T. Cognitive pretest and posttest: Scoring and rationale for formal operations assessment. Unpublished manuscript, University of California at Riverside, undated.

6. Cox, J. R., & Griggs, R. A. The effect of prior experience in Wason's selection task. Manuscript in preparation, University of Florida, 1981.

REFERENCES

Arlin, P.K. Cognitive development in adulthood. A fifth stage? Developmental Psychology, 1975, 11, 602-606.

Arlin, P.K. Piagetian operations in problem finding. Developmental Psychology, 1977, 13, 297-298.

Barratt, B.B. Training and transfer in combinatorial problem solving: The development of formal reasoning during early adolescence. Developmental Psychology, 1975, 11, 700-704.

Bart, W.M. A generalization of Piaget's logical mathematical model for the stage of formal operations. Journal of Mathematical Psychology, 1971, 8, 539-553.

Bart, W.M. A comparison of premise types in hypothetico-deductive thinking at the stage of formal operations. The Journal of Psychology, 1972, 81, 45-51.

Berzonsky, M.D., Weiner, A.S., & Raphael, D. Interdependence of formal reasoning. Developmental Psychology, 1975, 11, 258.

Blasi, A., & Hoeffel, E.C. Adolescence and formal operations. Human Development, 1974, 17, 344-363.

Bourne, L.E. Knowing and using concepts. Psychological Review, 1970, 77, 546-556.

Bourne, L.E. An inference model for conceptual rule learning. In R.L. Solso (Ed.), Theories in cognitive psychology: The Loyola Symposium. Hillsdale, NJ: Erlbaum, 1974.

Bracewell, R.J., & Hidi, S.E. The solution of an inferential problem as a function of stimulus materials. Quarterly Journal of Experimental Psychology, 1974, 26, 480-488.

Braine, M. On the relation between the natural logic of reasoning and standard logic. Psychological Review, 1978, 85, 1-21.

Brainerd, C.J. Piaget's theory of intelligence. Englewood Cliffs, NJ: Prentice-Hall, 1978.

Broughton, J.M. Piaget's structural developmental psychology II. Logic and psychology. Human Development, 1981, 24, 195-224.

Brown, C., Keats, J.A., Keats, D.M., & Seggie, I. Reasoning about implication: A comparison about Malaysian and Australian subjects. Journal of Cross-Cultural Psychology, 1980, 11, 395-410.

Bruner, J.S., Goodnow, J.J., & Austin, G.A. A study of thinking. New York: Wiley, 1956.

Burt, C. The development of reasoning in school children. Journal of Experimental Pedagogy, 1919, 5, (263), 68-77, 121-127.

Bynum, T.W., Thomas, J.A., & Weitz, L.J. Truth-functional logic in formal operational thinking: Inhelder & Piaget's evidence. Developmental Psychology, 1972, 7, 129-132.

Capon, N., & Kuhn, D. Logical reasoning in the supermarket: Adult females' use of a proportional reasoning strategy in an everyday context. Developmental Psychology, 1979, 15, 450-452.

Case, R. Intellectual development from birth to adulthood: A neo-Piagetian interpretation. In R.S. Siegler (Ed.), Children's thinking: What develops? Hillsdale, NJ: Erlbaum, 1978.

Cordell, R.L. Mature reasoning and problem solving. Unpublished master's thesis, University of Nottingham, 1978.

Cropper, D.A., Meck, D.S., & Ash, M.J. The relation between formal operations and a possible fifth stage of cognitive development. Developmental Psychology, 1977, 13, 517-518.

Danner, F.W., & Day, M.C. Eliciting formal operations. Child Development, 1977, 48, 1600-1606.

Dawes, R.M. An interpretation of Inhelder and Piaget. In Proceedings: Third invitational interdisciplinary seminar on Piagetian theory and its implications for the helping professions. Los Angeles: UAP, Childrens Hospital of Los Angeles, 1973.

Douglas, J.D., & Wong, A.C. Formal operations: Age and sex differences in Chinese and American children. Child Development, 1977, 48, 689-692.

Ennis, R.H. Children's ability to handle Piaget's propositional logic: A conceptual critique. Review of Educational Research, 1975, 45, 1-41.

Evans, J.St. B.T. Interpretation and "matching bias" in a reasoning task. Quarterly Journal of Experimental Psychology, 1972, 24, 193-199.

Evans, J. St. B.T. The psychology of deductive reasoning: Logic. In A. Burton & J. Radford (Eds.), Thinking in perspective. London: Methuen, 1978.

Evans, J. St. B.T. The psychology of deductive reasoning. London: Routledge and Kegan Paul, in press.

Evans, J. St. B.T., & Lynch, J.S. Matching bias in the selection task. British Journal of Psychology, 1973, 64, 391-397.

Fakouri, M.E. "Cognitive development in adulthood: A fifth stage?": A critique. Developmental Psychology, 1976, 12, 472.

Flavell, J.H. The developmental psychology of Jean Piaget. New York: D. Van Nostrand, 1963.

Flexer, B.K. & Roberge, J.J. IQ, field dependence-independence, and the development of formal operational thought. The Journal of General Psychology, 1980, 103, 191-201.

Ginsburg, H., & Opper, S. Piaget's theory of intellectual development (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall, 1969.

Griffiths, D.E. Physics teaching: Does it hinder intellectual development? American Journal of Physics, 1976, 44, 81-85.

Griggs, R.A., & Cox, J.R. The elusive thematic-materials effect in Wason's selection task. British Journal of Psychology, 1981, in press.

Helwig, J.T., & Council, K.A. (Eds.). SAS User's Guide: 1979 Edition. Raleigh, NC: SAS Institute, Inc., 1979.

Inhelder, B., & Piaget, J. The growth of logical thinking from childhood to adolescence. New York: Basic Books, 1958.

Johnson-Laird, P.N., & Tagart, J. How implication is understood. American Journal of Psychology, 1969, 2, 367-373.

Johnson-Laird, P.N., Legrenzi, P., & Sonino Legrenzi, M. Reasoning and a sense of reality. British Journal of Psychology, 1972, 63, 395-400.

Keating, D.P. Precocious cognitive development at the level of formal operations. Child Development, 1975, 46, 276-280.

Keating, D.P. Thinking processes in adolescence. In J. Abelson (Ed.), Handbook of adolescent psychology. New York: Wiley, 1980.

Keating, D.P., & Clark, L.V. Development of physical and social reasoning in adolescence. Developmental Psychology, 1980, 16, 23-30.

Keating, D.P., & Schaefer, R.A. Ability and sex differences in the acquisition of formal operations. Developmental Psychology, 1975, 11, 531-532.

Kleinbaum, D.G., & Kupper, L.L. Applied regression analyses and other multivariable methods. North Scituate, MA: Duxbury Press, 1978.

Kuhn, D., & Angelev, J. An experimental study of the development of formal operational thought. Child Development, 1976, 47, 697-706.

Kuhn, D., & Brannock, J. Development of the isolation of variables scheme in experimental and "natural experiment" contexts. Developmental Psychology, 1977, 13, 9-14.

Kuhn, D., Ho, V., & Adams, C. Formal reasoning among pre- and late adolescents. Child Development, 1979, 50, 1128-1135.

Lawson, A.E., & Blake, A.J.D. Concrete and formal thinking abilities in high school biology students as measured by three separate instruments. Journal of Research in Science Teaching, 1976, 13, 227-235.

Lawson, A.E., Karplus, R., & Adi, H. The acquisition of propositional logic and formal operational schemata during the secondary school years. Journal of Research in Science Teaching, 1978, 15, 465-478.

Lawson, A.E., & Nordland, F.H. The factor structure of some Piagetian tasks. _Journal of Research in Science Teaching_, 1976, _13_, 461-466.

Lawson, A.E., & Renner, J.W. A quantitative analysis of responses to Piagetian tasks and its implications for curriculum. _Science Education_, 1974, _58_, 545-559.

Linn, M.C. Cognitive style, training and formal thought. _Child Development_, 1978, _49_, 874-877.

Lovell, K. A follow-up study of Inhelder and Piaget's _The growth of logical thinking_. _British Journal of Psychology_, 1961, _52_, 143-153.

Manktelow, K.I., & Evans, J. St. B.T. Facilitation of reasoning by realism: Effect or non-effect? _British Journal of Psychology_, 1979, _70_, 477-488.

Martorano, S.C. A developmental analysis of performance on Piaget's formal operations tasks. _Developmental Psychology_, 1977, _13_, 666-672.

McKinnon, J.W., & Renner, J.W. Are colleges concerned with intellectual development? _American Journal of Physics_, 1971, _39_, 1047-1052.

Moshman, D. Consolidation and stage formation in the emergence of formal operations. _Developmental Psychology_, 1977, _13_, 95-100.

Moshman, D. Development of formal hypothesis testing ability. _Developmental Psychology_, 1979, _15_, 104-112.

Neimark, E.D. Longitudinal development of formal operations thought. _Genetic Psychology Monographs_, 1975, _91_, 171-225.

Neimark, E.D., & Slotnick, N.S. Development of understanding of logical connectives. _Journal of Educational Psychology_, 1970, _61_, 451-460.

O'Brien, T.C., & Shapiro, B.J. The development of logical thinking in children. _American Educational Research Journal_, 1968, _5_, 531-542.

Oltman, P.K., Raskin, E., & Witkin, H.A. _Group Embedded Figures Test_. Palo Alto, CA: Consulting Psychologists Press, 1971.

Osherson, D. Logic and models of logical thinking. In R.J. Falmagne (Ed.), Reasoning: Representation and process. Hillsdale, NJ: Erlbaum, 1975.

Paris, S.G. Comprehension of language connectives and propositional logical relationships. Journal of Experimental Child Psychology, 1973, 16, 278-291.

Parsons, C. Inhelder and Piaget's The Growth of Logical Thinking II: A logician's viewpoint. British Journal of Psychology, 1960, 51, 75-84.

Pascual-Leone, J. A mathematical model for the transition rule in Piaget's developmental stages. Acta Psychologica, 1970, 32, 301-345.

Peel, E.A. The nature of adolescent judgment. New York. Wiley, 1971.

Piaget, J. Intellectual evolution from adolescence to adulthood. Human Development, 1972, 15, 1-12.

Pollard, P. The effect of thematic content on the 'Wason selection task.' Current Psychological Research, 1981, 1, in press.

Pulos, S.M., & Linn, M.C. Generality of the controlling variables scheme. Journal of Early Adolescence, in press.

Restaino, L.C.R. Individual differences in logical operations: An information-processing analysis. In G.I. Lubin, M.K. Poulsen, J.F. Magery & M. Soto-McAlister (Eds.), Piagetian theory and the helping professions: 7th annual conference. (Volume II). USC Press, 1978.

Riley, J.M. Qualitative cognition and juvenile delinquency. In M.K. Poulsen & G.I. Lubin (Eds.), Piagetian theory and the helping professions: 8th annual conference, (Volume II). USC Press, 1979.

Rips, C.J., & Marcus, S.L. Supposition and the analysis of conditional sentences. In M.A. Just & P.A. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, NJ: Erlbaum, 1977.

Roberge, J.J. A study of children's abilities to reason with basic principles of deductive reasoning. American Educational Research Journal, 1970, 7, 583-595.

Roberge, J.J. Some effects of negation on adults' conditional reasoning abilities. Psychological Reports, 1971, 29, 839-844.

Rumelhart, D.E. Schemata: The building blocks of cognition. In R. Spiro, B. Bruce & W. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale, NJ: Erlbaum, 1980.

Saarni, C.I. Piagetian operations and field independence as factors in children's problem-solving performance. Child Development, 1973, 44, 338-345.

Sabatini, P.P. Age and professional specialization in formal reasoning (Doctoral dissertation, Wayne State University, 1979). Dissertation Abstracts International, 1980, 40, 5434B. (University Microfilms No. 8010172)

Scardamalia, M. Information processing capacity and the problem of horizontal decalage. A demonstration using combinatorial reasoning tasks. Child Development, 1977, 48, 28-37.

Schwebel, M. Formal operations in 1st year college students. The Journal of Psychology, 1975, 91, 133-141.

Shapiro, B.J., & O'Brien, T.C. Logical thinking in children ages six through thirteen. Child Development, 1970, 41, 823-829.

Siegler, R.S. The origins of scientific reasoning. In R.S. Siegler (Ed.), Children's thinking: What develops? Hillsdale, NJ: Erlbaum, 1978.

Staudenmayer, H. Understanding conditional reasoning with meaningful propositions. In R.J. Falmagne (Ed.), Reasoning: Representation and process. Hillsdale, NJ: Erlbaum, 1975.

Staudenmayer, H., & Bourne, L.E. Learning to interpret conditional sentences: A developmental study. Developmental Psychology, 1977, 13, 616-623.

Stephens, B., McLaughlin, J.A., Miller, C.K., & Glass, G.V. Factorial structure of selected psycho-educational measures and Piagetian reasoning assessments. Developmental Psychology, 1972, 6, 343-348.

Tomlinson-Keasey, C. Formal operations in females from eleven to fifty-four years of age. Developmental Psychology, 1972, 6, 364.

Van Duyne, P.C. Realism and linguistic complexity in reasoning. British Journal of Psychology, 1974, 65, 59-67.

Wason, P.C. On the failure to eliminate hypotheses in a conceptual task. Quarterly Journal of Experimental Psychology, 1960, 12, 129-140.

Wason, P.C. Reasoning. In B. Foss (Ed.), New horizons in psychology. London. Penguin, 1966.

Wason, P.C. The theory of formal operations--A critique. In B.A. Geber (Ed.), Piaget and knowing: Studies in genetic epistemology. London: Routledge & Kegan Paul, 1977(a).

Wason, P.C. Self-contradictions. In P.N. Johnson-Laird & P.C. Wason (Eds.), Thinking: Readings in cognitive science. Cambridge: Cambridge University Press, 1977(b).

Wason, P. Hypothesis testing and reasoning. Unit 25, Block 4, Cognitive psychology. Milton Keynes: Open University Press, 1978.

Wason, P.C., & Brooks, P.G. THOG: The anatomy of a problem. Psychological Research, 1979, 41, 79-90.

Wason, P., & Johnson-Laird, P. Psychology of reasoning: Structure and content. Cambridge, MA. Harvard University Press, 1972.

Wason, P.C., & Shapiro, D. Natural and contrived experience in a reasoning problem. Quarterly Journal of Experimental Psychology, 1971, 23, 63-71.

Webb, R.A. Concrete and formal operations in very bright 6- to 11-year-olds. Human Development, 1974, 17, 292-300.

Wildman, T.M., & Fletcher, H.J. Developmental increases and decreases in solutions of conditional syllogism problems. Developmental Psychology, 1977, 13, 630-636.

Wollman, W. Proportional reasoning and the validity of Piaget's theory of formal reasoning. In M.K. Poulsen & G.I. Lubin, (Eds.), Piagetian theory and the helping professions 8th annual conference, (Volume II). USC Press, 1979.

Wyatt, K.B., & Geis, M.F. Level of formal thought and organizational memory strategies. Developmental Psychology, 1978, 14, 433-434.

Yachanin, S.A. Differential effects of thematic materials and order of rule presentation on a reasoning task. Unpublished master's thesis, Bowling Green State University, 1980.

BIOGRAPHICAL SKETCH

[illegible] A. Warner was born April 29, 1955, in Charlotte, North Carolina. She attended schools in Charlotte and in Jacksonville, Florida, and in June, 1973, was graduated from Englewood High School in Jacksonville. Ms. Warner enrolled at the University of Florida in September, 1973. She received a Bachelor of Arts degree in psychology in 1977 and a Master of Arts degree in experimental psychology in 1979. Ms. Warner's areas of research interest are human problem solving and reasoning. She has been elected to Phi Beta Kappa, Phi Kappa Phi, and Sigma Xi. Ms. Warner taught General Psychology from 1978-1980 and in September of 1980 received the American Psychological Association Division Two Graduate Student Teaching Award. She has taken a position as an assistant professor at the University of Arizona in Tucson beginning in August, 1981.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Richard A. Griggs
Richard A. Griggs, Chairperson
Associate Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Patricia H. Miller
Patricia H. Miller
Associate Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Scott A. Miller
Scott A. Miller
Associate Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Keith D. White
Keith D. White
Associate Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mary Lou Koran
Mary L. Koran
Professor of Foundations of Education

This dissertation was submitted to the Graduate Faculty of the Department of Psychology in the College of Liberal Arts and Sciences and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy

December 1981

Dean for Graduate Studies and Research